

(19) World Intellectual Property Organization  
International Bureau



(43) International Publication Date  
26 July 2007 (26.07.2007)

PCT

(10) International Publication Number  
**WO 2007/084926 A2**

(51) International Patent Classification: **Not classified**

(21) International Application Number:  
PCT/US2007/060646

(22) International Filing Date: 17 January 2007 (17.01.2007)

(25) Filing Language: English

(26) Publication Language: English

(30) Priority Data:

60/759,298	17 January 2006 (17.01.2006)	US
60/790,373	7 April 2006 (07.04.2006)	US
60/791,178	11 April 2006 (11.04.2006)	US
60/812,702	9 June 2006 (09.06.2006)	US
60/836,998	11 August 2006 (11.08.2006)	US
60/860,358	21 November 2006 (21.11.2006)	US

(71) Applicants (for all designated States except US): **BI-OLEX, INC.** [US/US]; 4819 Emperor Blvd., Bristol Place, Suite 400, Durham, NC 27703 (US). **COMMONWEALTH SCIENTIFIC AND INDUSTRIAL RESEARCH ORGANISATION** [AU/AU]; Limestone Avenue, Campbell, ACT 2612 (AU).

(72) Inventors; and

(75) Inventors/Applicants (for US only): **DICKEY, Lynn, F.** [US/US]; 111 Buckden Place, Cary, NC 27511 (US). **COX, Kevin, M.** [US/US]; 1916 Sunset Drive, Raleigh, NC 27608 (US). **PEELE, Charles, G.** [US/US]; 3005 Atando Court, Apex, NC 27502 (US). **WANG, Ming-bo**

[AU/AU]; 3 Kiewa Street, Kaleen, Canberra, Australian Capital Territory (ATC) 2617 (AU).

(74) Agents: **HENRY, Leslie, T et al.; ALSTON & BIRD LLP**, Bank Of America Plaza, 101 South Tryon Street, Suite 4000, Charlotte, NC 28280-4000 (US).

(81) Designated States (unless otherwise indicated, for every kind of national protection available): AE, AG, AL, AM, AT, AU, AZ, BA, BB, BG, BR, BW, BY, BZ, CA, CH, CN, CO, CR, CU, CZ, DE, DK, DM, DZ, EC, EE, EG, ES, FI, GB, GD, GE, GH, GM, GT, HN, HR, HU, ID, IL, IN, IS, JP, KE, KG, KM, KN, KP, KR, KZ, LA, LC, LK, LR, LS, LT, LU, LV, LY, MA, MD, MG, MK, MN, MW, MX, MY, MZ, NA, NG, NI, NO, NZ, OM, PG, PH, PL, PT, RO, RS, RU, SC, SD, SE, SG, SK, SL, SM, SV, SY, TJ, TM, TN, TR, TT, TZ, UA, UG, US, UZ, VC, VN, ZA, ZM, ZW.

(84) Designated States (unless otherwise indicated, for every kind of regional protection available): ARIPO (BW, GH, GM, KE, LS, MW, MZ, NA, SD, SL, SZ, TZ, UG, ZM, ZW), Eurasian (AM, AZ, BY, KG, KZ, MD, RU, TJ, TM), European (AT, BE, BG, CH, CY, CZ, DE, DK, EE, ES, FI, FR, GB, GR, HU, IE, IS, IT, LT, LU, LV, MC, NL, PL, PT, RO, SE, SI, SK, TR), OAPI (BF, BJ, CF, CG, CI, CM, GA, GN, GQ, GW, ML, MR, NE, SN, TD, TG).

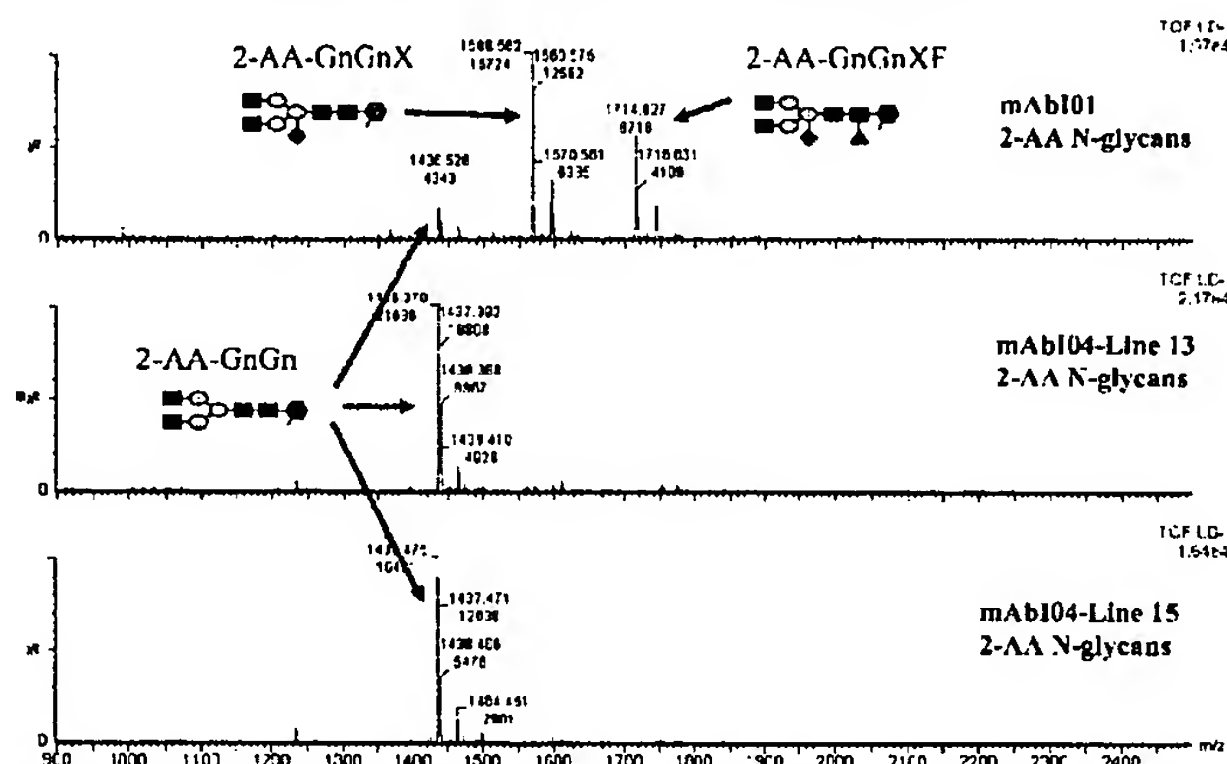
Published:

— without international search report and to be republished upon receipt of that report

[Continued on next page]

(54) Title: COMPOSITIONS AND METHODS FOR HUMANIZATION AND OPTIMIZATION OF N-GLYCANS IN PLANTS

MALDI analysis of mAb101 and mAb104 N-glycans



(57) Abstract: Methods for altering the N-glycosylation pattern of proteins in higher plants are provided. The methods comprise introducing into the plant a recombinant construct that provides for the inhibition of expression of  $\alpha$ 1,3-fucosyltransferase (FucT) and  $\beta$ 1,2-xylosyltransferase (XyIT) in a plant. Use of these constructs to inhibit or suppress expression of both of these enzymes, and isoforms thereof, advantageously provides for the production of endogenous and heterologous proteins having a "humanized" N-glycosylation pattern without impacting plant growth and development. Stably transformed higher plants having this protein N-glycosylation pattern are provided. Glycoprotein compositions, including monoclonal antibody compositions, having substantially homogeneous glycosylation profiles, and which are substantially homogeneous for the GO glycoform, are also provided.



---

*For two-letter codes and other abbreviations, refer to the "Guidance Notes on Codes and Abbreviations" appearing at the beginning of each regular issue of the PCT Gazette.*

## COMPOSITIONS AND METHODS FOR HUMANIZATION AND OPTIMIZATION OF *N*-GLYCANS IN PLANTS

### FIELD OF THE INVENTION

The present invention is related to the field of plant molecular biology, more particularly to the field of recombinant mammalian protein production in plants.

5

### BACKGROUND OF THE INVENTION

A number of plant species have been targeted for use in “molecular farming” of mammalian proteins of pharmaceutical interest. These plant expression systems provide for low cost production of biologically active mammalian proteins and are readily amenable to rapid and economical scale-up (Ma *et al.* (2003) *Nat. Rev. Genet.* 4:794-805; Raskin *et al.* (2002) *Trends Biotechnol.* 20:522-531). Large numbers of mammalian and plant proteins require post-translational processing for proper folding, assembly, and function. Of these modifications, the differences in glycosylation patterns between plants and mammals offer a challenge to the feasibility of plant expression systems to produce high quality recombinant mammalian proteins for pharmaceutical use.

15

As peptides move through the endoplasmic reticulum (ER) and Golgi subcellular compartments, sugar residue chains, or glycans, are attached, ultimately leading to the formation of glycoproteins. The linkage between the sugar chains and the peptide occurs by formation of a chemical bond to only one of four protein amino acids: asparagine, serine, threonine, and hydroxyllysine. Based on this linkage pattern, two basic types of sugar residue chains in glycoproteins have been recognized: the *N*-glycoside-linked sugar chain (also referred to as *N*-linked glycan or *N*-glycan), which binds to asparagine residues on the peptide; and the *O*-glycoside-linked sugar chain, which binds to serine, threonine, and hydroxylysine residues on the peptide.

20

The *N*-glycoside-linked sugar chains, or *N*-glycans, have various structures (see, for example, Takahashi, ed. (1989) *Biochemical Experimentation Method* 23 –

25

*Method for Studying Glycoprotein Sugar Chain* (Gakujutsu Shuppan Center), but share a common oligomannosidic core (see Figure 29A). The initial steps in the glycosylation pathway leading to the formation of *N*-glycans are conserved in plants and animals. However, the final steps involved in complex *N*-glycan formation differ (Lerouge *et al.* (1998) *Plant Mo. Biol.* 38:31-48; Steinkellner and Strasser (2003) *Ann. Plant Rev.* 9:181-192). Plants produce glycoproteins with complex *N*-glycans having a core bearing two N-acetylglucosamine (GlcNAc) residues that is similar to that observed in mammals. However, in plant glycoproteins this core is substituted by a  $\beta$ 1,2-linked xylose residue (core xylose), which residue does not occur in humans, Lewis<sup>a</sup> epitopes, and an  $\alpha$ 1,3-linked fucose (core  $\alpha$ [1,3]-fucose) instead of an  $\alpha$ 1,6-linked core fucose as in mammals (see, for example, Lerouge *et al.* (1998) *Plant Mol. Biol.* 38:31-48 for a review) (see also Figure 29B). Both the  $\alpha$ (1,3)-fucose and  $\beta$ (1,2)-xylose residues reportedly are, at least partly, responsible for the immunogenicity of plant glycoproteins in mammals (see, for example, Ree *et al.* (2000) *J. Biol. Chem.* 275:11451-11458; Bardor *et al.* (2003) *Glycobiol.* 13:427-434; Garcia-Casado *et al.* (1996) *Glycobiol.* 6:471-477). Therefore removal of these potentially allergenic sugar residues from mammalian glycoproteins recombinantly produced in plants would overcome concerns about the use of these proteins as pharmaceuticals for treatment of humans.

A number of recombinantly produced glycoproteins currently serve as therapeutics or are under clinical investigation. Examples include the interferons (IFNs), erythropoietin (EPO), tissue plasminogen activator (tPA), antithrombin, granulocyte-macrophage colony stimulating factor (GM-CSF), and therapeutic monoclonal antibodies (mABs). The oligosaccharide component of the *N*-glycan structures of glycoproteins can influence their therapeutic efficacy, as well as their physical stability, resistance to protease attack, pharmacokinetics, interaction with the immune system, and specific biological activity. See, for example, Jenkins *et al.* (1996) *Nature Biotechnol.* 14:975-981.

Methods are needed to alter the glycosylation pattern in plant expression systems, specifically to inhibit plant-specific glycosylation of the eukaryotic core structure, to advantageously produce recombinant mammalian proteins with a humanized glycosylation pattern.



## BRIEF SUMMARY OF THE INVENTION

Methods for altering the *N*-glycosylation pattern of proteins in higher plants are provided. The methods comprise stably transforming the plant with at least one recombinant nucleotide construct that provides for the inhibition of expression of  $\alpha$ 1,3-fucosyltransferase (FucT) and  $\beta$ 1,2-xylosyltransferase (XylT) in a plant. Use of these constructs to inhibit or suppress expression of one or both of these enzymes, and isoforms thereof, advantageously provides for the production of endogenous and heterologous proteins having a "humanized" *N*-glycosylation pattern without impacting plant growth and development. Stably transformed higher plants having this protein *N*-glycosylation pattern are provided. In some embodiments, the plant is a crop plant that is a member of the dicots, such as pea, alfalfa, and tobacco; in other embodiments, the plant is a crop plant that is a monocot, such as rice or maize. In yet other embodiments, the plant is a member of the *Lemnaceae* family, for example, a *Lemna* sp.

The transgenic plants of the invention have the ability to produce recombinant mammalian proteins having an *N*-glycosylation pattern that is more similar to that of the mammalian host. Thus, in some embodiments, the recombinantly produced mammalian proteins are glycoproteins comprising complex *N*-glycans having a reduction in the attachment of the plant-specific  $\alpha$ (1,3)-fucose and  $\beta$ (1,2)-xylose residues. In other embodiments, these recombinantly produced glycoproteins comprise complex *N*-glycans that are devoid of these plant-specific residues. In yet other embodiments, these recombinantly produced glycoproteins have GlcNAc<sub>2</sub>Man<sub>3</sub>GlcNAc<sub>2</sub> as the single glycan species attached to the asparagine glycosylation site(s) within the glycoprotein.

In some embodiments, the recombinantly produced glycoprotein is a monoclonal antibody. In this manner, the present invention provides transgenic plants that are capable of producing glycan-optimized monoclonal antibodies that specifically bind a target protein of interest, including monoclonal antibodies having increased effector function. Thus, in some embodiments, the recombinantly produced glycan-optimized monoclonal antibodies comprise complex *N*-glycans that have a reduction in the attachment of  $\alpha$ (1,3)-linked fucose residues, thereby increasing ADCC activity of these antibodies. In other embodiments, the recombinantly produced monoclonal antibodies comprise complex *N*-linked glycans that are devoid

of these plant-specific fucose residues. In this manner, the present invention provides for the production of a monoclonal antibody composition, wherein at least 90% or more of the intact antibody is represented by a single glycoform, more particularly, the G0 glycoform. Thus, in some embodiments of the invention, the recombinantly produced monoclonal antibodies have increased effector function, wherein the ADCC activity is increased and/or the ratio of ADCC/CDC activity is increased. In some of these embodiments, the recombinantly produced monoclonal antibodies have decreased CDC activity, which can advantageously reduce the potential for adverse side effects related to CDC activation upon administration. These glycan-optimized monoclonal antibodies advantageously can be used to alter current routes of administration and current therapeutic regimens, as their increased effector function means they can be dosed at lower concentrations and with less frequency, thereby reducing the potential for antibody toxicity and/or development of antibody tolerance. Furthermore, their improved effector function yields new approaches to treating clinical indications that have been previously been resistant or refractory to treatment with the corresponding monoclonal antibody produced in other recombinant host systems.

Compositions for practicing the methods of the invention are provided. The compositions comprise novel isolated polynucleotides and polypeptides encoding a *Lemna minor*  $\alpha$ 1,3-fucosyltransferase and  $\beta$ 1,2-xylosyltransferase, and variants and fragments thereof. Recombinant nucleotide constructs that target expression of these two proteins, or expression of variants thereof, are also provided, as are plant cells, plant tissues, plants, and seeds comprising these recombinant constructs.

## BRIEF DESCRIPTION OF THE FORMAL DRAWINGS

Figure 1 sets forth the DNA (SEQ ID NO:1; coding sequence set forth in SEQ ID NO:2) and amino acid (SEQ ID NO:3) sequences for the *Lemna minor*  $\alpha$ 1,3-fucosyltransferase (FucT). The coding sequence is shown in bold. Nucleotides denoted by the single underline (—) correspond to the FucT forward fragment within the RNAi expression cassette designed to inhibit expression of FucT (see Figure 5); nucleotides denoted by the double underline (==) correspond to the spacer sequence within this RNAi expression cassette. The FucT reverse fragment of the RNAi expression cassette is the antisense of the FucT forward fragment shown here.

Figure 2 sets forth an alignment of the *Lemna minor* FucT of SEQ ID NO:3 with  $\alpha$ 1,3-fucosyltransferases from other higher plants.

Figure 3 sets forth the DNA (SEQ ID NO:4; coding sequence set forth in SEQ ID NO:5) sequence for the *Lemna minor*  $\beta$ 1,2-xylosyltransferase (XylT) isoform #1 and the encoded amino acid (SEQ ID NO:6) sequence. Nucleotides denoted by the single underline (—) correspond to the XylT forward fragment within the RNAi expression cassette designed to inhibit expression of XylT (see Figure 6); nucleotides denoted by the double underline (==) correspond to the spacer sequence within this RNAi expression cassette. The XylT reverse fragment of the RNAi expression cassette is the antisense of the XylT forward fragment shown here.

Figure 4 sets forth an alignment of the *Lemna minor* XylT of SEQ ID NO:6 with  $\beta$ 1,2-xylosyltransferases from other higher plants.

Figure 5 sets forth one strategy for designing a single-gene RNAi knockout of *Lemna minor* FucT.

Figure 6 sets forth one strategy for designing a single-gene RNAi knockout of *Lemna minor* XylT based on the DNA sequence for XylT isoform #1.

Figure 7 sets forth one strategy for designing a double-gene RNAi knockout of *Lemna minor* FucT and XylT where the XylT portion of the RNAi knockout is based on the DNA sequence for XylT isoform #1.

Figure 8 shows the Fuc02 construct comprising an RNAi expression cassette designed for single-gene RNAi knockout of *Lemna minor* FucT. Expression of the FucT inhibitory sequence (denoted by FucT forward and FucT reverse arrows; see Figure 5) is driven by an operably linked expression control element (denoted as AocsAocsAocsAmasPmas) comprising three upstream activating sequences (Aocs) derived from the *Agrobacterium tumefaciens* octopine synthase gene operably linked to a promoter derived from an *Agrobacterium tumefaciens* mannopine synthase gene (AmasPmas). RbcS leader, rubisco small subunit leader sequence; ADH1, intron of maize alcohol dehydrogenase 1 gene; nos-ter, *Agrobacterium tumefaciens* nopaline synthetase (nos) terminator sequence.

Figure 9 shows the Xyl02 construct comprising an RNAi expression cassette designed for single-gene RNAi knockout of *Lemna minor* XylT. Expression of the XylT inhibitory sequence (denoted by XylT forward and XylT reverse arrows; see Figure 6) is driven by the operably linked AocsAocsAocsAmasPmas expression

control element. RbcS leader, rubisco small subunit leader sequence; ADH1, intron of maize alcohol dehydrogenase 1 gene; nos-ter, *Agrobacterium tumefaciens* nopaline synthetase (nos) terminator sequence.

Figure 10 shows the XF02 construct comprising a chimeric RNAi expression cassette designed for double-gene RNAi knockout of *Lemna minor* FucT/XylT. The hairpin RNA is expressed as a chimeric sequence (a chimeric hairpin RNA), where fragments of the two genes are fused together and expressed as one transcript. Expression of the FucT/XylT inhibitory sequence (denoted by FucT and XylT forward arrows and XylT and FucT reverse arrows; see Figure 7) is driven by the operably linked AocsAocsAocsAmasPmas expression control element. RbcS leader, rubisco small subunit leader sequence; ADH1, intron of maize alcohol dehydrogenase 1 gene; nos-ter, *Agrobacterium tumefaciens* nopaline synthetase (nos) terminator sequence.

Figure 11 shows the XF03 construct comprising an RNAi expression cassette designed for double-gene RNAi knockout of *Lemna minor* FucT/XylT. The cassette expresses two RNAi hairpins, one targeting expression of FucT, the other targeting expression of XylT. Expression of the FucT inhibitory sequence (denoted by FucT forward and FucT reverse arrows; see Figure 5) is driven by an operably linked expression control element comprising the *Lemna minor* ubiquitin promoter plus 5' UTR (LmUbq promoter) and intron (LmUbq intron) (see SEQ ID NO:7). Expression of the XylT inhibitory sequence (denoted by XylT forward and XylT reverse arrows; see Figure 6) is driven by the operably linked AocsAocsAocsAmasPmas expression control element. RbcS leader, rubisco small subunit leader sequence; ADH1, intron of maize alcohol dehydrogenase 1 gene; nos-ter, *Agrobacterium tumefaciens* nopaline synthetase (nos) terminator sequence.

Figure 12 shows the mAbI04 construct that provides for co-expression of an IgG1 monoclonal antibody (referred to herein as mAbI) and the double-gene knockout of *Lemna minor* FucT and XylT, wherein a chimeric hairpin RNA targeting expression of the FucT and XylT is expressed. Expression of the FucT/XylT inhibitory sequence (denoted by FucT and XylT forward arrows and XylT and FucT reverse arrows; see Figure 7) is driven by an operably linked expression control element comprising the *Spirodella polyrrhiza* ubiquitin promoter plus 5' UTR (SpUbq promoter) and intron (SpUbq intron) (see SEQ ID NO:8). Expression of the IgG1

light chain is driven by an operably linked expression control element comprising the *L. minor* ubiquitin promoter plus 5' UTR (LmUbq promoter) and intron (LmUbq intron). Expression of the IgG1 heavy chain is driven by the operably linked AocsAocsAocsAmasPmas expression control element. RbcS leader, rubisco small subunit leader sequence; ADH1, intron of maize alcohol dehydrogenase 1 gene; nos-ter, *Agrobacterium tumefaciens* nopaline synthetase (nos) terminator sequence.

Figure 13 shows the mAbI05 construct that provides for co-expression of mAbI and the double-knockout of *Lemna minor* FucT and XylT, wherein two hairpin RNAs are expressed, one targeting expression of FucT, the other targeting expression of the XylT. Expression of the FucT inhibitory sequence (denoted by FucT forward and FucT reverse arrows; see Figure 5) is driven by an operably linked expression control element comprising the *S. polyrrhiza* ubiquitin promoter plus 5' UTR (SpUbq promoter) and intron (SpUbq intron). Expression of the XylT inhibitory sequence (denoted by XylT forward and XylT reverse arrows; see Figure 6) is driven by an operably linked expression control element comprising the *Lemna aequinoctialis* ubiquitin promoter plus 5' UTR (LaUbq promoter) and intron (LaUbq intron) (see SEQ ID NO:9). Expression of the IgG1 light chain is driven by an operably linked expression control element comprising the *L. minor* ubiquitin promoter plus 5' UTR (LmUbq promoter) and intron (LmUbq intron). Expression of the IgG1 heavy chain is driven by the operably linked AocsAocsAocsAmasPmas expression control element. RbcS leader, rubisco small subunit leader sequence; ADH1, intron of maize alcohol dehydrogenase 1 gene; nos-ter, *Agrobacterium tumefaciens* nopaline synthetase (nos) terminator sequence.

Figure 14 shows the mAbI01 construct that provides for expression of mAbI, where FucT and XylT expression are not suppressed. Expression of the IgG1 light chain and IgG1 heavy chain are independently driven by the operably linked AocsAocsAocsAmasPmas expression control element. RbcS leader, rubisco small subunit leader sequence; ADH1, intron of maize alcohol dehydrogenase 1 gene; nos-ter, *Agrobacterium tumefaciens* nopaline synthetase (nos) terminator sequence. mAbI01 is referred to as the "wild-type" mAbI01 construct as the expressed mAbI exhibits the glycosylation profile of wild-type *L. minor*.

Figures 15 and 16 show the primary screening data for transgenic RNAi *L. minor* plant lines comprising the XF02 construct of Figure 10.



Figure 17 shows primary screening data for transgenic RNAi *L. minor* plant lines comprising the mAbI04 construct of Figure 12 and mAbI05 construct of Figure 13.

Figure 18 shows the structure and molecular weight of derivatized wild-type *L. minor* mAb *N*-glycans. “GnGn” represents the GlcNAc<sub>2</sub>Man<sub>3</sub>GlcNAc<sub>2</sub> *N*-glycan species, also referred to as a G0 *N*-glycan species. “GnGnX” represents the GlcNAc<sub>2</sub>Man<sub>3</sub>GlcNAc<sub>2</sub> *N*-glycan species with the plant-specific  $\beta$ (1,2)-xylose residue attached. “GnGnXF” represents the GlcNAc<sub>2</sub>Man<sub>3</sub>GlcNAc<sub>2</sub> *N*-glycan species with the plant-specific  $\beta$ (1,2)-xylose residue and plant-specific  $\alpha$ (1,3)-fucose residue attached.

Figure 19 shows that the wild-type mAbI01 construct (shown in Figure 15) providing for expression of the mAbI monoclonal IgG1 antibody in *L. minor*, without RNAi suppression of *L. minor* FucT and XylT, produces an *N*-glycosylation profile with three major *N*-glycan species, including one species having the  $\beta$ 1,2-linked xylose and one species having both the  $\beta$ 1,2-linked xylose and core  $\alpha$ 1,3-linked fucose residues; this profile is confirmed with liquid chromatography mass spectrometry (LC-MS) (Figure 20) and MALDI (Figure 21) analysis.

Figure 22 shows an overlay of the relative amounts of the various *N*-glycan species of mAbI produced in the wild-type *L. minor* line comprising the mAbI01 construct (no suppression of FucT or XylT) and in the two transgenic *L. minor* lines comprising the mAbI04 construct of Figure 12 (providing for coexpression of mAbI and the chimeric RNAi construct targeting both *L. minor* FucT and XIT). Note the enrichment of the GnGn (i.e., G0) glycan species, with no  $\beta$ 1,2-linked xylose or core  $\alpha$ 1,3-linked fucose residues attached, and the absence of the species having the  $\beta$ 1,2-linked xylose or both the  $\beta$ 1,2-linked xylose and core  $\alpha$ 1,3-linked fucose residues. “MGn” represents an *N*-glycan precursor, wherein the trimannose core structure, Man<sub>3</sub>GlcNAc<sub>2</sub>, has one N-acetylglucosamine attached to the 1,3 mannose arm. “GnM” represents an *N*-glycan precursor, wherein the trimannose core structure, Man<sub>3</sub>GlcNAc<sub>2</sub>, has one N-acetylglucosamine attached to the 1,6 mannose arm. These *N*-glycan precursors represent a trace amount of the total *N*-glycans present in the sample. This profile is confirmed with mass spec (LC-MS) (Figure 23) and MALDI (Figure 24) analysis.

Figure 25 shows intact mass analysis of mAbI compositions produced in wild-type *L. minor* (line 20) comprising the mAbI01 construct. When XylT and FucT expression are not suppressed in *L. minor*, the recombinantly produced mAbI composition is heterogeneous, comprising at least 9 different glycoforms, with the G0XF<sup>3</sup> glycoform being the predominate species present. Note the very minor peak representing the G0 glycoform.

Figure 26 shows intact mass analysis of mAbI compositions produced in transgenic *L. minor* (line 15) comprising the mAbI04 construct of Figure 12. When XylT and FucT expression are suppressed in *L. minor* using this chimeric RNAi construct, the intact mAbI composition is substantially homogeneous for G0 *N*-glycans, with only trace amounts of precursor *N*-glycans present (represented by the GnM and MGn precursor glycan species). In addition, the mAbI composition is substantially homogeneous for the G0 glycoform, wherein both glycosylation sites are occupied by the G0 *N*-glycan species, with three minor peaks reflecting trace amounts of precursor glycoforms (one peak showing mAbI having an Fc region wherein the C<sub>H</sub>2 domain of one heavy chain has a G0 glycan species attached to Asn 297, and the C<sub>H</sub>2 domain of the other heavy chain is unglycosylated; another peak showing mAbI having an Fc region wherein the C<sub>H</sub>2 domain of one heavy chain has a G0 glycan species attached to Asn 297, and the C<sub>H</sub>2 domain of the other heavy chain has the GnM or MGn precursor glycan attached to Asn 297; and another peak showing mAbI having an Fc region wherein the Asn 297 glycosylation site on each of the C<sub>H</sub>2 domains has a G0 glycan species attached, with a third G0 glycan species attached to an additional glycosylation site within the mAbI structure).

Figure 27 shows intact mass analysis of the mAbI compositions produced in transgenic *L. minor* (line 72) comprising the mAbI05 construct of Figure 13. When XylT and FucT expression are suppressed in *L. minor* using this construct, the intact mAbI composition is substantially homogeneous for G0 *N*-glycans, with only trace amounts of precursor *N*-glycan species present (represented by the GnM and MGn precursor glycan species). In addition, the mAbI composition is substantially homogeneous (at least 90%) for the G0 glycoform, with the same three minor peaks reflecting precursor glycoforms as obtained with the mAbI04 construct.

Figure 28 summarizes two possible designs for targeting expression of individual FucT and XylT genes.



Figure 29A shows the common oligomannosidic core structure of complex N-glycans of glycoproteins produced in plants and animals. In mammals, the core structure can include a fucose residue in which 1-position of the fucose is bound to 6-position of the N-acetylglucosamine in the reducing end through an  $\alpha$  bond (i.e.,  $\alpha(1,6)$ -linked fucose). Figure 29B shows the plant-specific modifications to these N-glycans. The mammalian R groups can be one of the following: (a)  $R=\text{GlcNAc}\beta(1,2)$ ; (b)  $R=\text{Gal}\beta(1,4)\text{-GlcNAc}\beta(1,2)$ ; (c)  $R=\text{NeuAc}\alpha(2,3)\text{-Gal}\beta(1,4)\text{-GlcNAc}\beta(1,2)$ ; (d)  $R=\text{NeuGc}\alpha(2,3)\text{-Gal}\beta(1,4)\text{-GlcNAc}\beta(1,2)$ ; and (e)  $R=\text{Gal}\alpha(1,3)\text{-Gal}\beta(1,4)\text{-GlcNAc}\beta(1,2)$ . The plant R groups can be one of the following: (a)  $R=\text{null}$ ; (b)  $R=\text{GlcNAc}\beta(1,2)$ ; (c),  $R=\text{Gal}\beta(1,3)\text{-GlcNAc}\beta(1,2)$  and (d)  $R=\text{Gal}\beta(1,4)\text{-GlcNAc}\beta(1,2)$ .



Abbreviations: Man, mannose; GlcNAc, N-acetylglucosamine; Xyl, xylose; Fuc, fucose; Gal, galactose; NeuAc (neuraminic acid (sialic acid)); \*, reducing end of sugar chain that binds to asparagine.

Figure 30 shows the G0, G0X, and G0XF<sup>3</sup> species of N-linked glycans of glycoproteins referred to in the description and claims of the present invention, along with the alternate nomenclature used herein.

Figure 31 sets forth the partial cDNA (SEQ ID NO:19; coding sequence set forth in SEQ ID NO:20) sequence for the *Lemna minor*  $\beta 1,2$ -xylosyltransferase (XylT) isoform #2, and partial amino acid (SEQ ID NO:21) sequence encoded thereby. Nucleotides denoted by the single underline (—) correspond to the XylT forward fragment within the RNAi expression cassette designed to inhibit expression of XylT (see Figure 33); nucleotides denoted by the double underline (==) correspond to the spacer sequence within this RNAi expression cassette. The XylT reverse fragment of the RNAi expression cassette is the antisense of the XylT forward fragment shown here.

Figure 32 sets forth an alignment of the *Lemna minor* XylT isoform #1 of SEQ ID NO:6 with the *Lemna minor* partial-length XylT isoform #2 of SEQ ID NO:21.

Figure 33 sets forth one strategy for designing a single-gene RNAi knockout of *Lemna minor* XylT based on the partial DNA sequence for XylT isoform #2.

Figure 34 sets forth one strategy for designing a double-gene RNAi knockout of *Lemna minor* FucT and XylT, where the XylT portion of the RNAi knockout is based on the partial DNA sequence for XylT isoform #2.

Figure 35 shows receptor binding activity of the CHO-derived and SP2/0-derived mAbI product for the FcγRIIIa on freshly isolated human NK cells.

Figure 36 shows receptor binding activity of the wild-type *Lemna*-derived mAbI product and the transgenic *Lemna*-derived mAbI product for the FcγRIIIa on freshly isolated human NK cells collected from Donor 1.

Figure 37 shows receptor binding activity of the wild-type *Lemna*-derived mAbI product and the transgenic *Lemna*-derived mAbI product for the FcγRIIIa on freshly isolated human NK cells collected from Donors 2 and 3.

Figure 38 shows receptor binding activity of the Sp2/0-derived mAbI product, the wild-type *Lemna*-derived mAbI product, and the transgenic *Lemna*-derived mAbI product for the mouse FcγRIV.

Figure 39 shows a diagram of the MDXA04 binary expression vector for RNAi silencing of FucT and XylT activity in *Lemna*. Hatched regions show the position of the heavy (H) and light (L) chain variable region gene sequences of fully human mAbI kappa antibody MDX-060 and the chimeric hairpin RNA (RNAi) designed to target silencing of endogenous *Lemna* genes encoding FucT and XylT. Promoters: P1, P2, and P3; terminator: T; selectable marker: SM; left border: LB; right border, RB. The MDXA01 expression vector used to express the MDX-060 mAb in wild-type *Lemna* did not contain the hairpin RNA region.

Figure 40 shows glycosyltransferase activity in *Lemna* wild-type and MDX-060 LEX<sup>Opt</sup> RNAi lines. Microsomal membranes from wild-type (WT) and MDX-060 LEX<sup>Opt</sup> RNAi (line numbers are indicated) plants were incubated in the presence of a reaction buffer containing GDP-Fuc, UDP-Xyl and GnGn-dabsyl-peptide acceptor. Mass peaks corresponding to fucosylated (white bars) or xylosylated (black bars) products synthesized by microsomes from each line were measured by positive reflectron mode MALDI-TOF MS and normalized, in percent, to the WT positive control. Boiled wildtype membranes (BWT) indicate background ion counts.

Figure 41 shows SDS-PAGE of plant extracts and protein A or hydroxyapatite purified samples from MDX-060 LEX<sup>Opt</sup> under non-reducing (Figure 41A) and reducing (Figure 41B) conditions, respectively. MAb purified from a CHO cell line

(MDX-060 CHO) was used as a positive control. Mark12 molecular weight markers were included on the gels. Gels were stained with Colloidal Blue.

Figure 42 shows the spectra obtained from negative, reflectron mode MALDI-TOF mass spectrometric analysis of 2-AA labeled *N*-glycans released from MDX-060 mAbs expressed in CHO (MDX-060 CHO), wild-type *Lemna* (MDX-060 LEX), or *Lemna* transformed with the XylT/FucT RNAi construct (MDX-060 LEX<sup>Opt</sup>). Significant peaks are identified by the corresponding mass ( $[M-H]^-$ ). The \* indicates the location of matrix artefacts.

Figure 43 shows the spectra obtained from NP-HPLC-QTOF MS analysis of 2-AA labeled *N*-glycans released from MDX-060 mAbs expressed in CHO (MDX-060 CHO), wild-type *Lemna* (MDX-060 LEX), or *Lemna* transformed with the XylT/FucT RNAi construct (MDX-060 LEX<sup>Opt</sup>). 2-AA labeled *N*-glycans were separated by normal phase chromatography and detected by fluorescence. The most abundant peaks from each sample (labeled a-i) were characterized by on-line negative mode QTOF MS and their corresponding QTOF mass spectra ( $[M-2H]^{2-}$ ) are shown.

Figure 44 shows *in vitro* activity of MDX-060 mAbs as measured by flow cytometric analysis of MDX-060 CHO, LEX, or glyco-optimized LEX<sup>Opt</sup> mAb binding to CD30 expressed on L540 cells. L540 cells were incubated with increasing concentrations of the indicated antibody as outlined in Example 6 herein below. Geo Mean Fluorescence Intensity (GMFI) is plotted against the various concentrations of mAb used. ■: MDX-060 CHO; ▲: MDX-060 LEX; ▼: MDX-060 LEX<sup>Opt</sup>.

Figure 45 shows equilibrium binding of glyco-optimized and wild-type mAb to two different human FcγRIIIa allotypes (Val<sup>158</sup> or Phe<sup>158</sup>). The binding signal as a function of FcγRIIIa was fit to a one-site binding model. ■: MDX-060 CHO; ▲: MDX-060 LEX; ▼: MDX-060 LEX<sup>Opt</sup>.

Figure 46 shows ADCC activity of MDX-060 mAb derived from CHO, LEX (wild-type *Lemna* glycosylation), or LEX<sup>Opt</sup> (RNAi transgenic *Lemna*). Human effector cells from a FcγRIIIaPhe<sup>158</sup> homozygote donor and a FcγRIIIaPhe/Val<sup>158</sup> heterozygote donor were incubated with BATDA-labeled L540 cells at an effector:target ratio of 50:1 in the presence of increasing concentrations of the indicated antibodies. Specific percent lysis at each mAb concentration is plotted. Human mAb1 not recognizing antigen on L540 cells was used as an isotype control in all experiments. EC<sub>50</sub> values (μg/mL), binding constants and maximal percent lysis

were calculated using GraphPad Prism 3.0 software. ■: MDX-060 CHO; ▲: MDX-060 LEX; ▼: MDX-060 LEX<sup>Opt</sup>.

Figure 47 shows intact mass analysis of the MDX-060 LEX mAb compositions produced in wild-type *L. minor* comprising the MDXA01 construct. When XylT and FucT expression are not suppressed in *L. minor*, the recombinantly produced MDX-060 LEX mAb composition comprises at least 7 different glycoforms, with the G0XF<sup>3</sup> glycoform being the predominate species present. Note the absence of a peak representing the G0 glycoform.

Figure 48 shows glycan mass analysis of the heavy chain of the MDX-060 LEX mAb produced in wild-type *L. minor* comprising the MDXA01 construct. When XylT and FucT expression are not suppressed in *L. minor*, the predominate *N*-glycan species present is G0XF<sup>3</sup>, with additional major peaks reflecting the G0X species. Note the minor presence of the G0 glycan species.

Figure 49 shows intact mass analysis of the MDX-060 LEX<sup>Opt</sup> mAb compositions produced in transgenic *L. minor* comprising the MDXA04 construct. When XylT and FucT expression are suppressed in *L. minor*, the intact mAb composition contains only G0 *N*-glycans. In addition, the composition is substantially homogeneous for the G0 glycoform (peak 2), wherein both glycosylation sites are occupied by the G0 *N*-glycan species, with two minor peaks reflecting trace amounts of precursor glycoforms (peak 1, showing mAb having an Fc region wherein the C<sub>H</sub>2 domain of one heavy chain has a G0 glycan species attached to Asn 297, and the C<sub>H</sub>2 domain of the other heavy chain is unglycosylated; and peak 3, showing mAb having an Fc region wherein the Asn 297 glycosylation site on each of the C<sub>H</sub>2 domains has a G0 glycan species attached, with a third G0 glycan species attached to an additional glycosylation site within the mAb structure).

Figure 50 shows glycan mass analysis of the heavy chain of the MDX-060 LEX<sup>Opt</sup> mAb produced in transgenic *L. minor* comprising the MDXA04 construct. When XylT and FucT expression are suppressed in *L. minor*, the only readily detectable *N*-glycan species attached to the Asn 297 glycosylation sites of the C<sub>H</sub>2 domains of the heavy chains is G0.

Figures 51A (MALDI analysis) and 51B (HPLC analysis) show that the homogeneous glycosylation profile exhibited by mAbI produced in transgenic *L. minor* (line 24) comprising the mAbI04 RNAi construct was consistently observed

with scaled-up production. This glycosylation profile was consistent over the 8-month period of continuous maintenance of the transgenic line via clonal expansion.

Figure 52 shows that suppression of FucT and XylT expression using the chimeric RNAi mAbI04 construct of Figure 12 results in endogenous glycoproteins having a homogeneous glycosylation pattern consistent with that observed for recombinant glycoproteins. For this figure, the  $\beta$ 1,2-linked xylose residue attached to the trimannose core structure is designated by the star symbol.

Figure 53 shows the structure of complex *N*-glycans described in Example 6 below. M = mannose; Gn = *N*-acetylglucosamine; A = galactose; X = xylose; F = fucose.

#### DETAILED DESCRIPTION OF THE INVENTION

The present inventions now will be described more fully hereinafter with reference to the accompanying drawings, in which some, but not all embodiments of the inventions are shown. Indeed, these inventions may be embodied in many different forms and should not be construed as limited to the embodiments set forth herein; rather, these embodiments are provided so that this disclosure will satisfy applicable legal requirements. Like numbers refer to like elements throughout.

Many modifications and other embodiments of the inventions set forth herein will come to mind to one skilled in the art to which these inventions pertain having the benefit of the teachings presented in the foregoing descriptions and the associated drawings. Therefore, it is to be understood that the inventions are not to be limited to the specific embodiments disclosed and that modifications and other embodiments are intended to be included within the scope of the appended claims. Although specific terms are employed herein, they are used in a generic and descriptive sense only and not for purposes of limitation.

The present invention provides compositions and methods for altering the *N*-glycosylation pattern of homologous and heterologous polypeptides produced in a plant, particularly a plant that serves as an expression system for recombinant proteins of interest. The methods comprise the use of nucleotide constructs comprising one or more sequences that are capable of inhibiting expression of  $\alpha$ 1,3-fucosyltransferase (FucT) and  $\beta$ 1,2-xylosyltransferase (XylT) in a plant. Compositions for use in practicing the methods of the invention, and compositions comprising glycoproteins

having a substantially homologous glycosylation profile, including glycan-optimized monoclonal antibodies with improved effector function, are also provided.

Definitions:

5           “Polypeptide” refers to any monomeric or multimeric protein or peptide.

          “Biologically active polypeptide” refers to a polypeptide that has the capability of performing one or more biological functions or a set of activities normally attributed to the polypeptide in a biological context. Those skilled in the art will appreciate that the term “biologically active” includes polypeptides in which the biological activity is altered as compared with the native protein (e.g., suppressed or enhanced), as long as the protein has sufficient activity to be of interest for use in industrial or chemical processes or as a therapeutic, vaccine, or diagnostics reagent. Biological activity can be determined by any method available in the art. For example, the biological activity of members of the interferon family of proteins can be determined by any of a number of methods including their interaction with interferon-specific antibodies, their ability to increase resistance to viral infection, or their ability to modulate the transcription of interferon-regulated gene targets. In like manner, biological activity of monoclonal antibodies can be determined by any of a number of methods including, but not limited to, assays for measuring binding specificity and effector function, for example, using assays for antibody-dependent cellular cytotoxicity (ADCC) and complement-dependent cytotoxicity (CDC) activity.

10  
15  
20

          By “host cell” is intended a cell that comprises a heterologous nucleic acid sequence of the invention. Though the nucleic acid sequences of the invention, and fragments and variants thereof, can be introduced into any cell of interest, of particular interest are plant host cells. In some embodiments, the plant host cells are cells of a plant that serves as a host for expression of recombinant proteins, for example, a plant expression system for production of recombinant mammalian proteins of interest as noted herein below.

25

          By “heterologous polypeptide of interest” is intended a polypeptide that is not expressed by the host cells in nature. Conversely, a “homologous polypeptide” is intended a polypeptide that is naturally produced within the cells of the host. Heterologous and homologous polypeptides that undergo post-translational *N*-glycosylation are referred to herein as heterologous or homologous glycoproteins. In

30



accordance with the methods of the present invention, the *N*-glycosylation pattern of both heterologous and homologous glycoproteins is altered within the cells of a plant host so that these glycoproteins have an *N*-glycosylation pattern that is more similar to that observed with mammalian hosts.

5 Exemplary heterologous polypeptide and heterologous nucleotide sequences of interest include, but are not limited to, sequences that encode mammalian polypeptides, such as insulin, growth hormone,  $\alpha$ -interferon,  $\beta$ -interferon,  $\beta$ -glucocerebrosidase,  $\beta$ -glucuronidase, retinoblastoma protein, p53 protein, angiostatin, leptin, erythropoietin (EPO), granulocyte macrophage colony stimulating factor,  
10 plasminogen, tissue plasminogen activator, blood coagulation factors, for example, Factor VII, Factor VIII, Factor IX, and activated protein C, alpha 1-antitrypsin, monoclonal antibodies (mAbs), Fab fragments, single-chain antibodies, cytokines, receptors, hormones, human vaccines, animal vaccines, peptides, and serum albumin.

For purposes of the present invention, the terms "*N*-glycan," "*N*-linked  
15 glycan," and "glycan" are used interchangeably and refer to an *N*-linked oligosaccharide, e.g., one that is or was attached by an N-acetylglucosamine (GlcNAc) residue linked to the amide nitrogen of an asparagine residue in a protein. The predominant sugars found on glycoproteins are glucose, galactose, mannose, fucose, N-acetylgalactosamine (GalNAc), N-acetylglucosamine (GlcNAc), and sialic  
20 acid (e.g., N-acetyl-neuraminic acid (NeuAc)). The processing of the sugar groups occurs cotranslationally in the lumen of the ER and continues in the Golgi apparatus for *N*-linked glycoproteins.

By "oligomannosidic core structure" or "trimannose core structure" of a complex *N*-glycan is intended the core structure shown in Figure 29A, wherein the  
25 core comprises three mannose (Man) and two N-acetylglucosamine (GlcNAc) monosaccharide residues that are attached to the asparagine residue of the glycoprotein. The asparagine residue is generally within the conserved peptide sequence Asn-Xxx-Thr or Asn-Xxx-Ser, where Xxx is any residue except proline, aspartate, or glutamate. Subsequent glycosylation steps yield the final complex *N*-  
30 glycan structure. The trimannose core structure is denoted herein as "Man<sub>3</sub>GlcNAc<sub>2</sub>."

The *N*-glycans attached to glycoproteins differ with respect to the number of branches (antennae) comprising peripheral sugars (e.g., GlcNAc, galactose, fucose, and sialic acid) that are added to the trimannose core structure. *N*-glycans are



commonly classified according to their branched constituents (e.g., complex, high mannose, or hybrid). A “complex” type *N*-glycan typically has at least one GlcNAc attached to the 1,3 mannose arm and at least one GlcNAc attached to the 1,6 mannose arm of a “trimannose” core. Where one GlcNAc is attached to each mannose arm, the species of *N*-linked glycan is denoted herein as “GlcNAc<sub>2</sub>Man<sub>3</sub>GlcNAc<sub>2</sub>” or “GnGn.” Where only one GlcNAc is attached, the *N*-glycan species is denoted herein as “GlcNAc<sub>1</sub>Man<sub>3</sub>GlcNAc<sub>2</sub>”, wherein the GlcNAc is attached to either the 1,3 mannose arm (denoted “MGn” herein) or the 1,6 mannose arm (denoted “GnM” herein) (see Figure 30). Complex *N*-glycans may also have galactose (“Gal”) or N-acetylgalactosamine (“GalNAc”) sugar residues that are optionally modified with sialic acid or derivatives (e.g., “NeuAc,” where “Neu” refers to neuraminic acid and “Ac” refers to acetyl). Where a galactose sugar residue is attached to each GlcNAc on each mannose arm, the species of *N*-linked glycan is denoted herein as “Gal<sub>2</sub>GlcNAc<sub>2</sub>Man<sub>3</sub>GlcNAc<sub>2</sub>.” Complex *N*-glycans may also have intrachain substitutions comprising “bisecting” GlcNAc and core fucose (“Fuc”). Complex *N*-glycans may also have multiple antennae on the “trimannose core,” often referred to as “multiple antennary glycans.” A “high mannose” type *N*-glycan has five or more mannose residues. A “hybrid” *N*-glycan has at least one GlcNAc on the terminal of the 1,3 mannose arm of the trimannose core and zero or more mannoses on the 1,6 mannose arm of the trimannose core.

The terms “G0 glycan” and “G0 glycan structure” and “G0 glycan species” are used interchangeably and are intended to mean the complex *N*-linked glycan having the GlcNAc<sub>2</sub>Man<sub>3</sub>GlcNAc<sub>2</sub> structure, wherein no terminal sialic acids (NeuAcs) or terminal galactose (Gal) sugar residues are present. If a G0 glycan comprises a fucose (“Fuc”) residue attached to the trimannose core structure, it is referred to herein as a “G0F<sup>3</sup> glycan” (having the plant-specific α1,3-linked fucose residue) or “G0F<sup>6</sup> glycan” (having the mammalian α1,6-linked fucose residue). In plants, a G0 glycan comprising the plant-specific β1,2-linked xylose residue attached to the trimannose core structure is referred to herein as a “G0X glycan,” and a G0 glycan comprising both the plant-specific β1,2-linked xylose residue and plant-specific α1,3-linked fucose residue attached to the trimannose core structure is referred to herein as a “G0XF<sup>3</sup> glycan.”

The terms "G1 glycan" and "G1 glycan structure" and "G1 glycan species" are used interchangeably and are intended to mean the complex *N*-linked glycan having the GlcNAc<sub>2</sub>Man<sub>3</sub>GlcNAc<sub>2</sub> structure, wherein one terminal galactose (Gal) residue is attached to either the 1,3 mannose or 1,6 mannose arm, and no terminal sialic acids are present. The terms "G2 glycan" and "G2 glycan structure" and "G2 glycan species" are used interchangeably and are intended to mean the complex *N*-linked glycan having the GlcNAc<sub>2</sub>Man<sub>3</sub>GlcNAc<sub>2</sub> structure, wherein a terminal galactose (Gal) residue is attached to the 1,3 mannose arm and the 1,6 mannose arm, and no terminal sialic acids are present.

The term "glycoform" as used herein refers to a glycoprotein containing a particular carbohydrate structure or structures. Thus, for example, a "G0 glycoform" refers to a glycoprotein that comprises only G0 glycan species attached to its glycosylation sites. It is recognized that a glycoprotein having more than one glycosylation site can have the same glycan species attached to each glycosylation site, or can have different glycan species attached to different glycosylation sites. In this manner, different patterns of glycan attachment yield different glycoforms of a glycoprotein.

The term "glycosylation profile" is intended to mean the characteristic "fingerprint" of the representative *N*-glycan species that have been released from a glycoprotein composition or glycoprotein product, either enzymatically or chemically, and then analyzed for their carbohydrate structure, for example, using LC-HPLC, or MALDI-TOF MS, and the like. See, for example, the review in *Current Analytical Chemistry*, Vol. 1, No. 1 (2005), pp. 28-57; herein incorporated by reference in its entirety.

The terms "substantially homogeneous," "substantially uniform," and "substantial homogeneity" in the context of a glycosylation profile for a glycoprotein composition or glycoprotein product are used interchangeably and are intended to mean a glycosylation profile wherein at least 80%, at least 85%, at least 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, or at least 99% of the total *N*-glycan species within the profile are represented by one desired *N*-glycan species, with a trace amount of precursor *N*-glycan species appearing in the profile. By "trace amount" is intended that any given precursor *N*-glycan species that is present in the glycosylation profile is present at less than 5%, preferably less than 4%, less than 3%, less than 2%,

less than 1%, and even less than 0.5% or even less than 0.1% of the total amount of *N*-glycan species appearing in the profile. By “precursor” *N*-glycan species is intended an *N*-glycan species that is incompletely processed. Examples of precursor *N*-glycan species present in trace amounts in the glycoprotein compositions or glycoprotein products of the invention, and thus appearing in the glycosylation profiles thereof, are the Man3GlcNAc2, MGn (GlcNAc1Man3GlcNAc2 wherein GlcNAc1 is attached to the 1,3 mannose arm), and GnM (GlcNAc1Man3GlcNAc2 wherein GlcNAc1 is attached to the 1,6 mannose arm) precursor *N*-glycan species described above.

Thus, for example, where the desired *N*-glycan species within a glycoprotein product or composition is G0, a substantially homogeneous glycosylation profile for that product or composition would be one wherein at least 80%, at least 85%, at least 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, or at least 99% of the total amount of *N*-glycan species appearing in the glycosylation profile for the product or composition is represented by the G0 glycan species, with a trace amount of precursor *N*-glycan species appearing in the glycosylation profile. For such a composition, a representative precursor *N*-glycan species appearing in its glycosylation profile would be the Man3GlcNAc2, MGn (GlcNAc1Man3GlcNAc2 wherein GlcNAc1 is attached to the 1,3 mannose arm), and GnM (GlcNAc1Man3GlcNAc2 wherein GlcNAc1 is attached to the 1,6 mannose arm) precursor *N*-glycan species described above.

The terms “substantially homogeneous,” “substantially uniform,” and “substantial homogeneity” in the context of a glycoprotein composition or glycoprotein product are used interchangeably and are intended to mean the glycoprotein product or glycoprotein composition wherein at least 80%, at least 85%, at least 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, or at least 99% of the glycoprotein present in the product or composition is represented by one desired glycoform, with a trace amount of precursor or undesired glycoforms being present in the composition. By “trace amount” is intended that any given precursor or undesired glycoform that is present in the glycoprotein product or composition is present at less than 5%, preferably less than 4%, less than 3%, less than 2%, less than 1%, and even less than 0.5% or even less than 0.1% of the total glycoprotein. By “precursor” glycoform is intended a glycoform wherein at least one glycosylation site is either unglycosylated, or is occupied by an *N*-glycan species that represents a precursor of

the desired *N*-glycan species, or a glycoform wherein one or more additional glycosylation sites is present, relative to the desired glycoform, and is occupied by (i.e., has attached thereto) the desired *N*-glycan species or an undesired *N*-glycan species.

5           Thus, for example, a substantially homogeneous glycoprotein composition or product comprising the G0 glycoform is a composition or product wherein at least 80%, 80%, at least 85%, at least 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, or at least 99% of the glycoprotein present in the product or composition is represented by the G0 glycoform, wherein all anticipated glycosylation sites are  
10           occupied by the G0 glycan species, with a trace amount of precursor or undesired glycoforms being present in the composition. In such a composition, a representative precursor glycoform would be one in which glycosylation sites are unoccupied, and an exemplary undesired glycoform would be a glycoform having a mixture of G0 glycan and G0X or G0XF3 glycan species attached to its glycosylation sites.

15           The term “antibody” is used in the broadest sense and covers fully assembled antibodies, antibody fragments that can bind antigen (e.g., Fab', F'(ab)<sub>2</sub>, Fv, single chain antibodies, diabodies), and recombinant peptides comprising the foregoing. Antibodies represent one of the many glycoproteins contemplated by the methods and compositions of the present invention. Derivatives of antibodies are also  
20           contemplated by the present invention. Derivatives include fusion proteins comprising an immunoglobulin or portion thereof, such as an Fc region having a C<sub>H</sub>2 domain.

          The term “monoclonal antibody” as used herein refers to an antibody obtained from a population of substantially homogeneous antibodies, i.e., the individual  
25           antibodies comprising the population are identical except for possible naturally occurring mutations that may be present in minor amounts.

          “Native antibodies” and “native immunoglobulins” are usually heterotetrameric glycoproteins of about 150,000 daltons, composed of two identical  
light (L) chains and two identical heavy (H) chains. Each light chain is linked to a  
30           heavy chain by one covalent disulfide bond, while the number of disulfide linkages varies among the heavy chains of different immunoglobulin isotypes. Each heavy and light chain also has regularly spaced intrachain disulfide bridges. Each heavy chain has at one end a variable domain (V<sub>H</sub>) followed by a number of constant domains.

Each light chain has a variable domain at one end ( $V_L$ ) and a constant domain at its other end; the constant domain of the light chain is aligned with the first constant domain of the heavy chain, and the light chain variable domain is aligned with the variable domain of the heavy chain. Particular amino acid residues are believed to form an interface between the light- and heavy-chain variable domains.

The term “variable” refers to the fact that certain portions of the variable domains differ extensively in sequence among antibodies and are used in the binding and specificity of each particular antibody for its particular antigen. However, the variability is not evenly distributed throughout the variable domains of antibodies. It is concentrated in three segments called complementarity determining regions (CDRs) or hypervariable regions both in the light-chain and the heavy-chain variable domains. The more highly conserved portions of variable domains are called the framework (FR) regions. The variable domains of native heavy and light chains each comprise four FR regions, largely adopting a  $\beta$ -sheet configuration, connected by three CDRs, which form loops connecting, and in some cases forming part of, the  $\beta$ -sheet structure. The CDRs in each chain are held together in close proximity by the FR regions and, with the CDRs from the other chain, contribute to the formation of the antigen-binding site of antibodies (see Kabat *et al.* (1991) *NIH Publ. No. 91-3242*, Vol. I, pages 647-669).

The constant domains are not involved directly in binding an antibody to an antigen, but exhibit various effector functions, such as Fc receptor (FcR) binding, participation of the antibody in antibody-dependent cellular cytotoxicity (ADCC), opsonization, initiation of complement-dependent cytotoxicity (CDC activity), and mast cell degranulation.

“Antibody fragments” comprise a portion of an intact antibody, preferably the antigen-binding or variable region of the intact antibody. Examples of antibody fragments include Fab, Fab', F(ab')<sub>2</sub>, and Fv fragments; diabodies; linear antibodies (Zapata *et al.* (1995) *Protein Eng.* 8(10):1057-1062); single-chain antibody molecules; and multispecific antibodies formed from antibody fragments. Papain digestion of antibodies produces two identical antigen-binding fragments, called “Fab” fragments, each with a single antigen-binding site, and a residual “Fc” fragment, whose name reflects its ability to crystallize readily. Pepsin treatment



yields an F(ab')<sub>2</sub> fragment that has two antigen-combining sites and is still capable of cross-linking antigen.

“Fv” is the minimum antibody fragment that contains a complete antigen recognition and binding site. In a two-chain Fv species, this region consists of a dimer of one heavy- and one light-chain variable domain in tight, non-covalent association. In a single-chain Fv species, one heavy- and one light-chain variable domain can be covalently linked by flexible peptide linker such that the light and heavy chains can associate in a “dimeric” structure analogous to that in a two-chain Fv species. It is in this configuration that the three CDRs of each variable domain interact to define an antigen-binding site on the surface of the V<sub>H</sub>-V<sub>L</sub> dimer. Collectively, the six CDRs confer antigen-binding specificity to the antibody. However, even a single variable domain (or half of an Fv comprising only three CDRs specific for an antigen) has the ability to recognize and bind antigen, although at a lower affinity than the entire binding site.

The Fab fragment also contains the constant domain of the light chain and the first constant domain (C<sub>H1</sub>) of the heavy chain. Fab fragments differ from Fab' fragments by the addition of a few residues at the carboxy terminus of the heavy-chain C<sub>H1</sub> domain including one or more cysteines from the antibody hinge region. Fab'-SH is the designation herein for Fab' in which the cysteine residue(s) of the constant domains bear a free thiol group. F(ab')<sub>2</sub> antibody fragments originally were produced as pairs of Fab' fragments that have hinge cysteines between them. Other chemical couplings of antibody fragments are also known.

The “light chains” of antibodies (immunoglobulins) from any vertebrate species can be assigned to one of two clearly distinct types, called kappa (κ) and lambda (λ), based on the amino acid sequences of their constant domains.

Depending on the amino acid sequence of the constant domain of their heavy chains, immunoglobulins can be assigned to different classes. There are five major classes of human immunoglobulins: IgA, IgD, IgE, IgG, and IgM, and several of these may be further divided into subclasses (isotypes), e.g., IgG1, IgG2, IgG3, IgG4, IgA, and IgA2. The heavy-chain constant domains that correspond to the different classes of immunoglobulins are called alpha, delta, epsilon, gamma, and mu, respectively. The subunit structures and three-dimensional configurations of different classes of immunoglobulins are well known. Different isotypes have different effector

functions. For example, human IgG1 and IgG3 isotypes mediate antibody-dependent cell-mediated cytotoxicity (ADCC) activity.

Immunoglobulins have conserved *N*-linked glycosylation of the Fc region of each of the two heavy chains. Thus, for example, immunoglobulins of the IgG type have glycosylated C<sub>H</sub>2 domains bearing *N*-linked oligosaccharides at asparagine 297 (Asn-297). Different glycoforms of immunoglobulins exist depending upon the particular *N*-glycan species attached to each of these two glycosylation sites, and depending upon the degree to which both sites are glycosylated within an immunoglobulin composition.

“Nucleotide sequence of interest” as used herein with reference to expression of heterologous polypeptides refers to any polynucleotide sequence encoding a heterologous polypeptide intended for expression in a host, particularly a plant host, for example, in a higher plant, including members of the dicotyledonaceae and monocotyledonaceae. For example, polynucleotide sequences encoding therapeutic (e.g., for veterinary or medical uses) or immunogenic (e.g., for vaccination) polypeptides can be expressed using transformed plant hosts, for example, duckweed, according to the present invention.

The use of the term “polynucleotide” is not intended to limit the present invention to polynucleotides comprising DNA. Those of ordinary skill in the art will recognize that polynucleotides can comprise ribonucleotides and combinations of ribonucleotides and deoxyribonucleotides. Such deoxyribonucleotides and ribonucleotides include both naturally occurring molecules and synthetic analogues. The polynucleotides of the invention also encompass all forms of sequences including, but not limited to, single-stranded forms, double-stranded forms, hairpins, stem-and-loop structures, and the like.

The terms “inhibit,” “inhibition,” and “inhibiting” as used herein refer to any decrease in the expression or function of a target gene product, including any relative decrement in expression or function up to and including complete abrogation of expression or function of the target gene product. The term “expression” as used herein in the context of a gene product refers to the biosynthesis of that gene product, including the transcription and/or translation and/or assembly of the gene product. Inhibition of expression or function of a target gene product (i.e., a gene product of interest) can be in the context of a comparison between any two plants, for example,



expression or function of a target gene product in a genetically altered plant versus the expression or function of that target gene product in a corresponding wild-type plant. Alternatively, inhibition of expression or function of the target gene product can be in the context of a comparison between plant cells, organelles, organs, tissues, or plant parts within the same plant or between plants, and includes comparisons between developmental or temporal stages within the same plant or between plants. Any method or composition that down-regulates expression of a target gene product, either at the level of transcription or translation, or down-regulates functional activity of the target gene product can be used to achieve inhibition of expression or function of the target gene product.

The term “inhibitory sequence” encompasses any polynucleotide or polypeptide sequence that is capable of inhibiting the expression of a target gene product, for example, at the level of transcription or translation, or which is capable of inhibiting the function of a target gene product. Examples of inhibitory sequences include, but are not limited to, full-length polynucleotide or polypeptide sequences, truncated polynucleotide or polypeptide sequences, fragments of polynucleotide or polypeptide sequences, variants of polynucleotide or polypeptide sequences, sense-oriented nucleotide sequences, antisense-oriented nucleotide sequences, the complement of a sense- or antisense-oriented nucleotide sequence, inverted regions of nucleotide sequences, hairpins of nucleotide sequences, double-stranded nucleotide sequences, single-stranded nucleotide sequences, combinations thereof, and the like. The term “polynucleotide sequence” includes sequences of RNA, DNA, chemically modified nucleic acids, nucleic acid analogs, combinations thereof, and the like.

It is recognized that inhibitory polynucleotides include nucleotide sequences that directly (i.e., do not require transcription) or indirectly (i.e., require transcription or transcription and translation) inhibit expression of a target gene product. For example, an inhibitory polynucleotide can comprise a nucleotide sequence that is a chemically synthesized or *in vitro*-produced small interfering RNA (siRNA) or micro RNA (miRNA) that, when introduced into a plant cell, tissue, or organ, would directly, though transiently, silence expression of the target gene product of interest. Alternatively, an inhibitory polynucleotide can comprise a nucleotide sequence that encodes an inhibitory nucleotide molecule that is designed to silence expression of the gene product of interest, such as sense-orientation RNA, antisense RNA, double-

stranded RNA (dsRNA), hairpin RNA (hpRNA), intron-containing hpRNA, catalytic RNA, miRNA, and the like. In yet other embodiments, the inhibitory polynucleotide can comprise a nucleotide sequence that encodes a mRNA, the translation of which yields a polypeptide that inhibits expression or function of the target gene product of interest. In this manner, where the inhibitory polynucleotide comprises a nucleotide sequence that encodes an inhibitory nucleotide molecule or a mRNA for a polypeptide, the encoding sequence is operably linked to a promoter that drives expression in a plant cell so that the encoded inhibitory nucleotide molecule or mRNA can be expressed.

Inhibitory sequences are designated herein by the name of the target gene product. Thus, for example, an “ $\alpha$ 1,3-fucosyltransferase (FucT) inhibitory sequence” would refer to an inhibitory sequence that is capable of inhibiting the expression of a FucT, for example, at the level of transcription and/or translation, or which is capable of inhibiting the function of a FucT. Similarly, a “ $\beta$ 1,2-xylosyltransferase (XylT) inhibitory sequence” would refer to an inhibitory sequence that is capable of inhibiting the expression of a XylT, at the level of transcription and/or translation, or which is capable of inhibiting the function of a XylT. When the phrase “capable of inhibiting” is used in the context of a polynucleotide inhibitory sequence, it is intended to mean that the inhibitory sequence itself exerts the inhibitory effect; or, where the inhibitory sequence encodes an inhibitory nucleotide molecule (for example, hairpin RNA, miRNA, or double-stranded RNA polynucleotides), or encodes an inhibitory polypeptide (i.e., a polypeptide that inhibits expression or function of the target gene product), following its transcription (for example, in the case of an inhibitory sequence encoding a hairpin RNA, miRNA, or double-stranded RNA polynucleotide) or its transcription and translation (in the case of an inhibitory sequence encoding an inhibitory polypeptide), the transcribed or translated product, respectively, exerts the inhibitory effect on the target gene product (i.e., inhibits expression or function of the target gene product).

The term “introducing” in the context of a polynucleotide, for example, a nucleotide construct of interest, is intended to mean presenting to the plant the polynucleotide in such a manner that the polynucleotide gains access to the interior of a cell of the plant. Where more than one polynucleotide is to be introduced, these polynucleotides can be assembled as part of a single nucleotide construct, or as

separate nucleotide constructs, and can be located on the same or different transformation vectors. Accordingly, these polynucleotides can be introduced into the host cell of interest in a single transformation event, in separate transformation events, or, for example, in plants, as part of a breeding protocol. The methods of the invention do not depend on a particular method for introducing one or more polynucleotides into a plant, only that the polynucleotide(s) gains access to the interior of at least one cell of the plant. Methods for introducing polynucleotides into plants are known in the art including, but not limited to, transient transformation methods, stable transformation methods, and virus-mediated methods.

“Transient transformation” in the context of a polynucleotide is intended to mean that a polynucleotide is introduced into the plant and does not integrate into the genome of the plant.

By “stably introducing” or “stably introduced” in the context of a polynucleotide introduced into a plant is intended the introduced polynucleotide is stably incorporated into the plant genome, and thus the plant is stably transformed with the polynucleotide.

“Stable transformation” or “stably transformed” is intended to mean that a polynucleotide, for example, a nucleotide construct described herein, introduced into a plant integrates into the genome of the plant and is capable of being inherited by the progeny thereof, more particularly, by the progeny of multiple successive generations. In some embodiments, successive generations include progeny produced vegetatively (i.e., asexual reproduction), for example, with clonal propagation. In other embodiments, successive generations include progeny produced via sexual reproduction. A higher plant host that is “stably transformed” with at least one nucleotide construct that is capable of inhibiting expression of a FucT and/or XylT as described herein refers to a higher plant host that has the nucleotide construct(s) integrated into its genome, and is capable producing progeny, either via asexual or sexual reproduction, that also comprise the inhibitory nucleotide construct(s) stably integrated into their genome, and hence the progeny will also exhibit the desired phenotype of having an altered *N*-glycosylation pattern characterized by a reduction in the attachment of  $\alpha$ 1,3-fucose and/or  $\beta$ 1,2-xylose residues to the *N*-glycans of homologous and heterologous glycoproteins produced therein.

As used herein, the term “plant” includes reference to whole plants, plant organs (*e.g.*, leaves, stems, roots, *etc.*), seeds, plant cells, and progeny of same. Parts of transgenic plants are to be understood within the scope of the invention to comprise, for example, plant cells, protoplasts, tissues, callus, embryos as well as  
 5 flowers, ovules, stems, fruits, leaves, roots, root tips, and the like originating in transgenic plants or their progeny previously transformed with a DNA molecule of the invention and therefore consisting at least in part of transgenic cells. As used herein, the term “plant cell” includes, without limitation, cells of seeds, embryos, meristematic regions, callus tissue, leaves, roots, shoots, gametophytes, sporophytes,  
 10 pollen, and microspores.

The class of plants that can be used in the methods of the invention is generally as broad as the class of higher plants amenable to transformation techniques, including both monocotyledonous (monocot) and dicotyledonous (dicot) plants. Examples of dicots include, but are not limited to, legumes including  
 15 soybeans and alfalfa, tobacco, potatoes, tomatoes, and the like. Examples of monocots include, but are not limited to, maize, rice, oats, barley, wheat, members of the duckweed family, grasses, and the like. “Lower-order plants” refers to non-flowering plants including ferns, horsetails, club mosses, mosses, liverworts, hornworts, algae, for example, red, brown, and green algae, gametophytes,  
 20 sporophytes of pteridophytes, and the like. In some embodiments, the plant of interest is a member of the duckweed family of plants.

The term “duckweed” refers to members of the family *Lemnaceae*. This family currently is divided into five genera and 38 species of duckweed as follows:  
 genus *Lemna* (*L. aequinoctialis*, *L. disperma*, *L. ecuadoriensis*, *L. gibba*, *L. japonica*,  
 25 *L. minor*, *L. miniscula*, *L. obscura*, *L. perpusilla*, *L. tenera*, *L. trisulca*, *L. turionifera*, *L. valdiviana*); genus *Spirodela* (*S. intermedia*, *S. polyrrhiza*, *S. punctata*); genus *Wolffia* (*Wa. angusta*, *Wa. arrhiza*, *Wa. australina*, *Wa. borealis*, *Wa. brasiliensis*, *Wa. columbiana*, *Wa. elongata*, *Wa. globosa*, *Wa. microscopica*, *Wa. neglecta*); genus *Wolffiella* (*Wl. caudata*, *Wl. denticulata*, *Wl. gladiata*, *Wl. hyalina*, *Wl. lingulata*, *Wl. repunda*, *Wl. rotunda*, and *Wl. neotropica*) and genus *Landoltia* (*L. punctata*). Any  
 30 other genera or species of *Lemnaceae*, if they exist, are also aspects of the present invention. *Lemna* species can be classified using the taxonomic scheme described by Landolt (1986) *Biosystematic Investigation on the Family of Duckweeds: The family*

of *Lemnaceae*—A Monograph Study (Geobotanischen Institut ETH, Stiftung Rubel, Zurich).

The term “duckweed nodule” as used herein refers to duckweed tissue comprising duckweed cells where at least about 50%, 55%, 60%, 65%, 70%, 75%, 80%, 85%, 90%, 95%, or 100% of the cells are differentiated cells. A “differentiated cell,” as used herein, is a cell with at least one phenotypic characteristic (e.g., a distinctive cell morphology or the expression of a marker nucleic acid or protein) that distinguishes it from undifferentiated cells or from cells found in other tissue types. The differentiated cells of the duckweed nodule culture described herein form a tiled smooth surface of interconnected cells fused at their adjacent cell walls, with nodules that have begun to organize into frond primordium scattered throughout the tissue. The surface of the tissue of the nodule culture has epidermal cells connected to each other via plasmadesmata. Members of the duckweed family reproduce by clonal propagation, and thus are representative of plants that clonally propagate.

“Duckweed-preferred codons” as used herein refers to codons that have a frequency of codon usage in duckweed of greater than 17%.

“*Lemna*-preferred codons” as used herein refers to codons that have a frequency of codon usage in the genus *Lemna* of greater than 17%.

“*Lemna gibba*-preferred codons” as used herein refers to codons that have a frequency of codon usage in *Lemna gibba* of greater than 17% where the frequency of codon usage in *Lemna gibba* was obtained from the Codon Usage Database (GenBank Release 113,0; at [http://www.kazusa.or.jp/codon/cgibin/showcodon.cgi?species=Lemna+gibba+\[gbpln\]](http://www.kazusa.or.jp/codon/cgibin/showcodon.cgi?species=Lemna+gibba+[gbpln])).

“Translation initiation codon” refers to the codon that initiates the translation of the mRNA transcribed from the nucleotide sequence of interest.

“Translation initiation context nucleotide sequence” as used herein refers to the identity of the three nucleotides directly 5' of the translation initiation codon.

“Secretion” as used herein refers to translocation of a polypeptide across both the plasma membrane and the cell wall of a host plant cell.

“Operably linked” as used herein in reference to nucleotide sequences refers to multiple nucleotide sequences that are placed in a functional relationship with each other. Generally, operably linked DNA sequences are contiguous and, where necessary to join two protein coding regions, in reading frame.



### Isolated Polynucleotides and Polypeptides

The present invention provides isolated polynucleotides and polypeptides that are involved in further modification of plant *N*-linked glycans (also referred to as “*N*-glycans”), particularly an  $\alpha$ 1,3-fucosyltransferase (FucT) and  $\beta$ 1,2-xylosyltransferase (XylT) identified in *Lemna minor*, a member of the duckweed family, and variants and fragments of these polynucleotides and polypeptides. Inhibition of the expression of one or both of these proteins, or biologically active variants thereof, in a plant that expresses these proteins beneficially yields an *N*-glycosylation pattern that has a reduction in the attachment of  $\alpha$ 1,3-fucose and  $\beta$ 1,2-xylose residues to glycoprotein *N*-glycans. In some embodiments of the invention, the methods disclosed herein provide for complete inhibition of expression of FucT and XylT, yielding an *N*-glycosylation pattern of glycoproteins produced within a plant wherein the *N*-linked glycans are devoid of  $\alpha$ 1,3-fucose and  $\beta$ 1,2-xylose residues.

The full-length cDNA sequence, including 5'- and 3'-UTR, for *L. minor* alpha 1-3 fucosyltransferase (FucT) is set forth in Figure 1; see also SEQ ID NO:1 (open reading frame set forth in SEQ ID NO:2). The predicted amino acid sequence encoded thereby is set forth in SEQ ID NO:3. At least two isoforms of the *L. minor* FucT gene have been identified; the homology between the isoforms is about 90%. The encoded protein shares some similarity with other FucTs from other higher plants. See Figure 2. For example, the *L. minor* FucT sequence shares approximately 50.1% sequence identity with the *Arabidopsis thaliana* FucT shown in Figure 2.

The full-length cDNA sequence, including 5'- and 3'-UTR, for *L. minor*  $\beta$ 1-2 xylosyltransferase (XylT) (isoform #1) is set forth in Figure 3; see also SEQ ID NO:4 (ORF set forth in SEQ ID NO:5). The predicted amino acid sequence encoded thereby is set forth in SEQ ID NO:6. At least two isoforms of the *L. minor* XylT gene have been identified; the homology between the isoforms is about 90%. The encoded protein shares some similarity with other XylTs from other higher plants. See Figure 4. For example, the *L. minor* XylT shares approximately 56.4% sequence identity with the *Arabidopsis thaliana* XylT shown in Figure 4. A partial-length cDNA sequence, including 3'-UTR, for *L. minor*  $\beta$ 1-2 xylosyltransferase (XylT) (isoform #2) is set forth in Figure 3I; see also SEQ ID NO:19 (ORF set forth in SEQ ID NO:20). The predicted amino acid sequence encoded thereby is set forth in SEQ ID NO:21.

The partial-length XylT isoform #2 shares high sequence identity with the corresponding region of the full-length XylT isoform #1, as can be seen from the alignment shown in Figure 32.

5 The invention encompasses isolated or substantially purified polynucleotide or protein compositions. An "isolated" or "purified" polynucleotide or protein, or biologically active portion thereof, is substantially or essentially free from components that normally accompany or interact with the polynucleotide or protein as found in its naturally occurring environment. Thus, an isolated or purified polynucleotide or protein is substantially free of other cellular material, or culture  
10 medium when produced by recombinant techniques, or substantially free of chemical precursors or other chemicals when chemically synthesized. Optimally, an "isolated" polynucleotide is free of sequences (optimally protein encoding sequences) that naturally flank the polynucleotide (i.e., sequences located at the 5' and 3' ends of the polynucleotide) in the genomic DNA of the organism from which the polynucleotide  
15 is derived. For example, in various embodiments, the isolated polynucleotide can contain less than about 5 kb, 4 kb, 3 kb, 2 kb, 1 kb, 0.5 kb, or 0.1 kb of nucleotide sequence that naturally flank the polynucleotide in genomic DNA of the cell from which the polynucleotide is derived. A protein that is substantially free of cellular material includes preparations of protein having less than about 30%, 20%, 10%, 5%,  
20 or 1% (by dry weight) of contaminating protein. When the protein of the invention or biologically active portion thereof is recombinantly produced, optimally culture medium represents less than about 30%, 20%, 10%, 5%, or 1% (by dry weight) of chemical precursors or non-protein-of-interest chemicals.

The coding sequence for the *L. minor* FucT gene is set forth as nucleotides (nt)  
25 243-1715 of SEQ ID NO:1 and as SEQ ID NO:2, and the amino acid sequence for the encoded FucT polypeptide is set forth in SEQ ID NO:3. The coding sequence for the *L. minor* XylT isoform #1 gene is set forth as nucleotides 63-1592 of SEQ ID NO:4 and as SEQ ID NO:5, and the amino acid sequence for the encoded XylT polypeptide is set forth in SEQ ID NO:6. The coding sequence for the partial-length *L. minor*  
30 XylT isoform #2 gene is set forth as nucleotides 1-1276 of SEQ ID NO:19 and as SEQ ID NO:20, and the amino acid sequence for the encoded partial-length XylT polypeptide is set forth in SEQ ID NO:21.



In particular, the present invention provides for isolated polynucleotides comprising nucleotide sequences encoding the amino acid sequences shown in SEQ ID NOS:3, 6, and 21. Further provided are polypeptides having an amino acid sequence encoded by a polynucleotide described herein, for example those set forth in  
5 SEQ ID NOS:1, 2, 4, 5, 19, and 20, and fragments and variants thereof. Nucleic acid molecules comprising the complements of these nucleotide sequences are also provided. It is recognized that the coding sequence for the FucT and/or XylT gene can be expressed in a plant for overexpression of the encoded FucT and/or XylT. However, for purposes of suppressing or inhibiting the expression of these proteins,  
10 the respective nucleotide sequences of SEQ ID NO:1, 2, 4, 5, 19, and 20 will be used to design constructs for suppression of expression of the respective FucT and/or XylT protein. Thus, polynucleotides, in the context of suppressing the FucT protein refers to the FucT coding sequences and to polynucleotides that when expressed suppress or inhibit expression of the FucT gene, for example, via direct or indirect suppression as  
15 noted herein below. Similarly, polynucleotides, in the context of suppressing or inhibiting the XylT protein refers to the XylT coding sequences and to polynucleotides that when expressed suppress or inhibit expression of the XylT gene, for example, via direct or indirect suppression as noted herein below.

Fragments and variants of the disclosed polynucleotides and proteins encoded  
20 thereby are also encompassed by the present invention. By "fragment" is intended a portion of the FucT or XylT polynucleotide or a portion of the FucT or XylT amino acid sequence encoded thereby. Fragments of a polynucleotide may encode protein fragments that retain the biological activity of the native protein and hence have FucT activity or XylT activity as noted elsewhere herein. Alternatively, fragments of a  
25 polynucleotide that are useful as hybridization probes generally do not encode fragment proteins retaining biological activity. Fragments of a FucT or XylT polynucleotide can also be used to design inhibitory sequences for suppression of expression of the FucT and/or XylT polypeptide. Thus, for example, fragments of a nucleotide sequence may range from at least about 15 nucleotides, 20 nucleotides,  
30 about 50 nucleotides, about 100 nucleotides, about 150 nucleotides, about 200 nucleotides, about 250 nucleotides, about 300 nucleotides, about 350 nucleotides, about 400 nucleotides, about 450 nucleotides, about 500 nucleotides, about 550 nucleotides, about 600 nucleotides, about 650 nucleotides, about 700 nucleotides,

about 750 nucleotides, about 800 nucleotides, and up to the full-length polynucleotide encoding the proteins of the invention.

A fragment of a FucT polynucleotide that encodes a biologically active portion of a FucT protein of the invention will encode at least 15, 25, 30, 50, 100,  
5 150, 200, 250, 300, 350, 400, 450, 475, 500 contiguous amino acids, or up to the total number of amino acids present in a full-length FucT protein of the invention (for example, 509 amino acids for SEQ ID NO:3). A fragment of a XylT polynucleotide that encodes a biologically active portion of a full-length XylT protein of the invention will encode at least 15, 25, 30, 50, 100, 150, 200, 250, 300, 350, 400, 450,  
10 475 contiguous amino acids, or up to the total number of amino acids present in a full-length XylT protein of the invention (for example, 490 amino acids for SEQ ID NO:3). A fragment of a XylT polynucleotide that encodes a biologically active portion of a partial-length XylT protein of the invention will encode at least 15, 25, 30, 50, 100, 150, 200, 250, 300, 350, 400 contiguous amino acids, or up to the total  
15 number of amino acids present in a partial-length XylT protein of the invention (for example, 490 amino acids for SEQ ID NO:21)

Thus, a fragment of a FucT or XylT polynucleotide may encode a biologically active portion of a FucT or XylT protein, respectively, or it may be a fragment that can be used as a hybridization probe or PCR primer, or used to design inhibitory  
20 sequences for suppression, using methods disclosed below. A biologically active portion of a FucT or XylT protein can be prepared by isolating a portion of one of the FucT or XylT polynucleotides of the invention, respectively, expressing the encoded portion of the FucT or XylT protein (e.g., by recombinant expression *in vitro*), and assessing the activity of the encoded portion of the FucT or XylT polypeptide.

Polynucleotides that are fragments of an FucT or XylT nucleotide sequence comprise  
25 at least 15, 20, 50, 75, 100, 150, 200, 250, 300, 350, 400, 450, 500, 550, 600, 650, 700, 800, 900, 1,000, 1,100, 1,200, 1,300, 1,400, or 1450 contiguous nucleotides, or up to the number of nucleotides present in a FucT or XylT polynucleotide disclosed herein (for example, 1865, 1473, 1860, 1530, 1282, or 1275 nucleotides for SEQ ID  
30 NOS:1, 2, 4, 5, 19, and 20, respectively).

“Variants” is intended to mean substantially similar sequences. For polynucleotides, a variant comprises a deletion and/or addition of one or more nucleotides at one or more sites within the native polynucleotide and/or a substitution

of one or more nucleotides at one or more sites in the native polynucleotide. As used herein, a “native” polynucleotide or polypeptide comprises a naturally occurring nucleotide sequence or amino acid sequence, respectively. For polynucleotides, conservative variants include those sequences that, because of the degeneracy of the genetic code, encode the amino acid sequence of one of the FucT or XylT polypeptides of the invention. Naturally occurring allelic variants such as these can be identified with the use of well-known molecular biology techniques, as, for example, with polymerase chain reaction (PCR) and hybridization techniques as outlined below. Variant polynucleotides also include synthetically derived polynucleotides, such as those generated, for example, by using site-directed mutagenesis but which still encode a FucT or XylT protein of the invention. Generally, variants of a particular polynucleotide of the invention (for example, SEQ ID NO:1, SEQ ID NO:2, SEQ ID NO:4, SEQ ID NO:5, SEQ ID NO:19, or SEQ ID NO:20, fragments thereof, and complements of these sequences) will have at least about 40%, 45%, 50%, 55%, 60%, 65%, 70%, 75%, 80%, 85%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, 99% or more sequence identity to that particular polynucleotide as determined by sequence alignment programs and parameters described elsewhere herein.

Variants of a particular polynucleotide of the invention (i.e., the reference polynucleotide) can also be evaluated by comparison of the percent sequence identity between the polypeptide encoded by a variant polynucleotide and the polypeptide encoded by the reference polynucleotide. Thus, for example, an isolated polynucleotide that encodes a polypeptide with a given percent sequence identity to the FucT or XylT polypeptide of SEQ ID NO:3, SEQ ID NO:6, or SEQ ID NO:21, respectively, is disclosed. Percent sequence identity between any two polypeptides can be calculated using sequence alignment programs and parameters described elsewhere herein. Where any given pair of polynucleotides of the invention is evaluated by comparison of the percent sequence identity shared by the two polypeptides they encode, the percent sequence identity between the two encoded polypeptides is at least about 40%, 45%, 50%, 55%, 60%, 65%, 70%, 75%, 80%, 85%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, 99% or more sequence identity.

“Variant” protein is intended to mean a protein derived from the native protein by deletion or addition of one or more amino acids at one or more sites in the native protein and/or substitution of one or more amino acids at one or more sites in the native protein. Variant proteins encompassed by the present invention are

5 biologically active, that is they continue to possess the desired biological activity of the native protein, that is, the enzymatic activity of attaching the  $\alpha$ 1,3-linked fucose residue (activity of FucT) or  $\beta$ 1,2-linked xylose residue (activity of XylT) to glycoprotein *N*-glycans in plants as described herein. Such variants may result from, for example, genetic polymorphism or from human manipulation. Biologically active

10 variants of a native FucT or XylT protein of the invention will have at least about 40%, 45%, 50%, 55%, 60%, 65%, 70%, 75%, 80%, 85%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, 99% or more sequence identity to the amino acid sequence for the native protein as determined by sequence alignment programs and parameters described elsewhere herein. A biologically active variant of a protein of

15 the invention may differ from that protein by as few as 1-15 amino acid residues, as few as 1-10, such as 6-10, as few as 5, as few as 4, 3, 2, or even 1 amino acid residue.

The proteins of the invention may be altered in various ways including amino acid substitutions, deletions, truncations, and insertions. Methods for such manipulations are generally known in the art. For example, amino acid sequence

20 variants and fragments of the FucT and XylT proteins can be prepared by mutations in the DNA. Methods for mutagenesis and polynucleotide alterations are well known in the art. See, for example, Kunkel (1985) *Proc. Natl. Acad. Sci. USA* 82:488-492; Kunkel *et al.* (1987) *Methods in Enzymol.* 154:367-382; U.S. Patent No. 4,873,192; Walker and Gaastra, eds. (1983) *Techniques in Molecular Biology* (MacMillan

25 Publishing Company, New York) and the references cited therein. Guidance as to appropriate amino acid substitutions that do not affect biological activity of the protein of interest may be found in the model of Dayhoff *et al.* (1978) *Atlas of Protein Sequence and Structure* (Natl. Biomed. Res. Found., Washington, D.C.), herein incorporated by reference. Conservative substitutions, such as exchanging one amino

30 acid with another having similar properties, may be optimal.

Thus, the polynucleotides of the invention include both the naturally occurring FucT and XylT sequences as well as mutant forms. Likewise, the proteins of the invention encompass both naturally occurring FucT and XylT proteins as well as

variations and modified forms thereof. Such variants will continue to possess the desired activity. Thus, where expression of a functional protein is desired, the expressed protein will possess the desired FucT or XylT activity, i.e., the enzymatic activity of attaching the  $\alpha$ 1,3-linked fucose residue (activity of FucT) or  $\beta$ 1,2-linked xylose residue (activity of XylT) to glycoprotein *N*-glycans in plants as described  
5 herein. Where the objective is inhibition of expression or function of the FucT and/or XylT polypeptide, the desired activity of the variant polynucleotide or polypeptide is one of inhibiting expression or function of the respective FucT and/or XylT polypeptide. Obviously, where expression of a functional FucT or XylT variant is  
10 desired, the mutations that will be made in the DNA encoding the variant must not place the sequence out of reading frame and optimally will not create complementary regions that could produce secondary mRNA structure. See, EP Patent Application Publication No. 75,444.

Where a functional protein is desired, the deletions, insertions, and  
15 substitutions of the protein sequences encompassed herein are not expected to produce radical changes in the characteristics of the protein. However, when it is difficult to predict the exact effect of the substitution, deletion, or insertion in advance of doing so, one skilled in the art will appreciate that the effect will be evaluated by routine screening assays, including the assays for monitoring FucT and XylT activity  
20 described herein below in the Experimental section.

Variant polynucleotides and proteins also encompass sequences and proteins derived from a mutagenic and recombinogenic procedure such as DNA shuffling. With such a procedure, one or more different FucT or XylT coding sequences can be manipulated to create a new FucT or XylT protein possessing the desired properties.  
25 In this manner, libraries of recombinant polynucleotides are generated from a population of related sequence polynucleotides comprising sequence regions that have substantial sequence identity and can be homologously recombined *in vitro* or *in vivo*. For example, using this approach, sequence motifs encoding a domain of interest may be shuffled between the FucT or XylT gene of the invention and other known FucT or  
30 XylT genes, respectively, to obtain a new gene coding for a protein with an improved property of interest. Strategies for such DNA shuffling are known in the art. See, for example, Stemmer (1994) *Proc. Natl. Acad. Sci. USA* 91:10747-10751; Stemmer (1994) *Nature* 370:389-391; Cramer *et al.* (1997) *Nature Biotech.* 15:436-438;



Moore *et al.* (1997) *J. Mol. Biol.* 272:336-347; Zhang *et al.* (1997) *Proc. Natl. Acad. Sci. USA* 94:4504-4509; Cramer *et al.* (1998) *Nature* 391:288-291; and U.S. Patent Nos. 5,605,793 and 5,837,458.

5       The comparison of sequences and determination of percent identity and percent similarity between two sequences can be accomplished using a mathematical algorithm. In a preferred embodiment, the percent identity between two amino acid sequences is determined using the Needleman and Wunsch (1970) *J. Mol. Biol.* 48:444-453 algorithm, which is incorporated into the GAP program in the GCG software package (available at [www.accelrys.com](http://www.accelrys.com)), using either a BLOSSUM62  
10       matrix or a PAM250 matrix, and a gap weight of 16, 14, 12, 10, 8, 6, or 4 and a length weight of 1, 2, 3, 4, 5, or 6. In yet another preferred embodiment, the percent identity between two nucleotide sequences is determined using the GAP program in the GCG software package, using a BLOSUM62 scoring matrix (see Henikoff *et al.* (1989) *Proc. Natl. Acad. Sci. USA* 89:10915) and a gap weight of 40, 50, 60, 70, or 80 and a  
15       length weight of 1, 2, 3, 4, 5, or 6. A particularly preferred set of parameters (and the one that should be used if the practitioner is uncertain about what parameters should be applied to determine if a molecule is within a sequence identity limitation of the invention) is using a BLOSUM62 scoring matrix with a gap weight of 60 and a length weight of 3).

20       The percent identity between two amino acid or nucleotide sequences can also be determined using the algorithm of Meyers and Miller (1989) *CABIOS* 4:11-17 which has been incorporated into the ALIGN program (version 2.0), using a PAM120 weight residue table, a gap length penalty of 12 and a gap penalty of 4.

      An alternative indication that two nucleic acid molecules are closely related is  
25       that the two molecules hybridize to each other under stringent conditions. Stringent conditions are sequence-dependent and are different under different environmental parameters. Generally, stringent conditions are selected to be about 5°C to 20°C lower than the thermal melting point ( $T_m$ ) for the specific sequence at a defined ionic strength and pH. The  $T_m$  is the temperature (under defined ionic strength and pH) at  
30       which 50% of the target sequence hybridizes to a perfectly matched probe. Conditions for nucleic acid hybridization and calculation of stringencies can be found, for example, in Sambrook *et al.* (2001) *Molecular Cloning: A Laboratory Manual* (Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY) and Tijssen (1993)



*Hybridization With Nucleic Acid Probes, Part I: Theory and Nucleic Acid Preparation* (Laboratory Techniques in Biochemistry and Molecular Biology, Elsevier Science Ltd., NY, NY).

5 For purposes of the present invention, "stringent conditions" encompass conditions under which hybridization will only occur if there is less than 25% mismatch between the hybridization molecule and the target sequence. "Stringent conditions" may be broken down into particular levels of stringency for more precise definition. Thus, as used herein, "moderate stringency" conditions are those under which molecules with more than 25% sequence mismatch will not hybridize;  
10 conditions of "medium stringency" are those under which molecules with more than 15% mismatch will not hybridize, and conditions of "high stringency" are those under which sequences with more than 10% mismatch will not hybridize. Conditions of "very high stringency" are those under which sequences with more than 6% mismatch will not hybridize.

15 The FucT and XylT polynucleotides of the invention can be used as probes for the isolation of corresponding homologous sequences in other organisms, more particularly in other plant species. In this manner, methods such as PCR, hybridization, and the like can be used to identify such sequences based on their sequence homology to the sequences of the invention. See, for example, Sambrook *et al.* (1989) *Molecular Cloning: A Laboratory Manual* (2d ed., Cold Spring Harbor  
20 Laboratory Press, Plainview, New York) and Innis *et al.* (1990), *PCR Protocols: A Guide to Methods and Applications* (Academic Press, New York). Polynucleotide sequences isolated based on their sequence identity to the entire FucT or XylT polynucleotides of the invention (i.e., SEQ ID NOS:1 and 2 for FucT; SEQ ID NOS:4 and 5 for XylT isoform #1 of SEQ ID NO:6; and SEQ ID NOS:19 and 20 for XylT  
25 isoform #2 of SEQ ID NO:21 ) or to fragments and variants thereof are encompassed by the present invention.

In a PCR method, oligonucleotides primers can be designed for use in PCR reactions for amplification of corresponding DNA sequences from cDNA or genomic  
30 DNA extracted from any organism of interest. Known methods of PCR include, but are not limited to, methods using paired primers, nested primers, single specific primers, degenerate primers, gene-specific primers, vector-specific primers, partially-mismatched primers, and the like. Methods for designing PCR primers and PCR

cloning are generally known in the art and are disclosed in Sambrook *et al.* (1989) *Molecular Cloning: A Laboratory Manual* (2d ed., Cold Spring Harbor Laboratory Press, Plainview, New York). See also Innis *et al.*, eds. (1990) *PCR Protocols: A Guide to Methods and Applications* (Academic Press, New York); Innis and Gelfand, eds. (1995) *PCR Strategies* (Academic Press, New York); and Innis and Gelfand, eds. (1999) *PCR Methods Manual* (Academic Press, New York).

In a hybridization method, all or part of a known nucleotide sequence can be used as a probe that selectively hybridizes to other corresponding polynucleotides present in a population of cloned genomic DNA fragments or cDNA fragments (i.e., cDNA or genomic libraries) from another organism of interest. The so-called hybridization probes may be genomic DNA fragments, cDNA fragments, RNA fragments, or other oligonucleotides, and may be labeled with a detectable group such as  $^{32}\text{P}$ , or any other detectable marker. Probes for hybridization can be made by labeling synthetic oligonucleotides based on the nucleotide sequence of interest, for example, the FucT or XylT polynucleotides of the invention. Degenerate primers designed on the basis of conserved nucleotides or amino acid residues in the known nucleotide or encoded amino acid sequence can additionally be used. Methods for construction of cDNA and genomic libraries, and for preparing hybridization probes, are generally known in the art and are disclosed in Sambrook *et al.* (1989) *Molecular Cloning: A Laboratory Manual* (2d ed., Cold Spring Harbor Laboratory Press, Plainview, New York), herein incorporated by reference.

For example, all or part of the specific known FucT or XylT polynucleotide sequence may be used as a probe that selectively hybridizes to other FucT or XylT nucleotide and messenger RNAs. To achieve specific hybridization under a variety of conditions, such probes include sequences that are unique and are preferably at least about 10 nucleotides in length, and more optimally at least about 20 nucleotides in length. This technique may be used to isolate other corresponding FucT or XylT nucleotide sequences from a desired organism or as a diagnostic assay to determine the presence of a FucT or XylT coding sequences in an organism. Hybridization techniques include hybridization screening of plated DNA libraries (either plaques or colonies; see, for example, Innis *et al.*, eds. (1990) *PCR Protocols: A Guide to Methods and Applications* (Academic Press, New York)).

Thus, in addition to the native FucT and XylT polynucleotides and fragments and variants thereof, the isolated polynucleotides of the invention also encompass homologous DNA sequences identified and isolated from other organisms by hybridization with entire or partial sequences obtained from the FucT or XylT polynucleotides of the invention or variants thereof. Conditions that will permit other DNA sequences to hybridize to the DNA sequences disclosed herein can be determined in accordance with techniques generally known in the art. For example, hybridization of such sequences may be carried out under various conditions of moderate, medium, high, or very high stringency as noted herein above.

#### Methods of the Invention

The present invention is directed to methods for altering protein glycosylation patterns in higher plants, particularly in higher plants that serve as hosts for production of recombinant proteins, particularly recombinant mammalian proteins of pharmaceutical interest. The methods find use in producing higher plants that are capable of producing recombinant proteins having an *N*-glycosylation pattern that more closely resembles that found in mammals. Compositions of the invention include higher plants that are stably transformed to comprise an altered *N*-glycosylation pattern of their endogenous (i.e., homologous) and recombinantly produced heterologous proteins. In some embodiments, the higher plants are transgenic plants that produce monoclonal antibodies (mAbs) to mammalian proteins that have enhanced ADCC activity relative to mAbs produced in a control plant that has not had the glycosylation machinery altered to reduce the plant-specific attachment of  $\alpha$ 1,3-fucose residues to the *N*-glycans of homologous and heterologous glycoproteins produced therein.

The methods of the invention target the suppression (i.e., inhibition) of the expression of one or both of the enzymes involved the production of complex glycoproteins in higher plants. Of particular interest is suppression of a fucosyltransferase or one or more isoforms thereof, suppression of a xylosyltransferase or one or more isoforms thereof, or suppression of expression of both of these proteins and one or more isoforms thereof. It is recognized that suppression of the fucosyltransferase and/or xylosyltransferase and one or more isoforms thereof can be accomplished transiently. Alternatively, by stably

suppressing the expression of the fucosyltransferase and/or xylosyltransferase, it is possible to produce transgenic higher plants that carry over from generation to generation, either asexually or sexually, the ability to produce glycoproteins having an *N*-glycosylation pattern that more closely resembles that found in mammals, more particularly, in humans. This advantageously provides for the production of recombinant mammalian glycoproteins that have reduced attachment of the plant  $\beta$ 1,2-linked xylose residue and/or  $\alpha$ 1,3-linked fucose residue to glycoprotein *N*-glycans.

Inhibition of the expression of one or advantageously both of these proteins in a plant, for example, a dicotyledonous or monocotyledonous plant, for example, a duckweed plant, can be carried out using any method known in the art. In this manner, a polynucleotide comprising an inhibitory sequence for FucT, XylT, or a combination thereof is introduced into the host cell of interest. For transient suppression, the FucT or XylT inhibitory sequence can be a chemically synthesized or *in vitro*-produced small interfering RNA (siRNA) or micro RNA (miRNA) that, when introduced into the host cell, would directly, though transiently, inhibit FucT, XylT, or a combination thereof, by silencing expression of these target gene product(s).

Alternatively, stable suppression of expression of FucT, XylT, or a combination thereof is desirable as noted herein above. Thus, in some embodiments, the activity of the FucT or the XylT polypeptide of the invention is reduced or eliminated by transforming a plant cell with an expression cassette that expresses a polynucleotide that inhibits the expression of the FucT or XylT, or both. The polynucleotide may inhibit the expression of the FucT or XylT, or both, directly, by preventing transcription or translation of the FucT or XylT messenger RNA, or indirectly, by encoding a polypeptide that inhibits the transcription or translation of a gene encoding the FucT or XylT, or both. Methods for inhibiting or eliminating the expression of a gene in a plant are well known in the art, and any such method may be used in the present invention to inhibit the expression of FucT or XylT, or both.

Thus, in some embodiments, expression of the FucT and/or XylT protein can be inhibited by introducing into the plant a nucleotide construct, such as an expression cassette, comprising a sequence that encodes an inhibitory nucleotide molecule that is designed to silence expression of the FucT and/or XylT gene product of interest, such as sense-orientation RNA, antisense RNA, double-stranded RNA (dsRNA), hairpin

RNA (hpRNA), intron-containing hpRNA, catalytic RNA, miRNA, and the like. In other embodiments, the nucleotide construct, for example, an expression cassette, can comprise a sequence that encodes a mRNA, the translation of which yields a polypeptide of interest that inhibits expression or function of the FucT and/or XylT gene product of interest. Where the nucleotide construct comprises a sequence that encodes an inhibitory nucleotide molecule or a mRNA for a polypeptide of interest, the sequence is operably linked to a promoter that drives expression in a plant cell so that the encoded inhibitory nucleotide molecule or mRNA can be expressed.

In accordance with the present invention, the expression of a FucT or XylT gene is inhibited if the protein level of the FucT or XylT is statistically lower than the protein level of the same FucT or XylT in a plant that has not been genetically modified or mutagenized to inhibit the expression of that FucT or XylT. In particular embodiments of the invention, the protein level of the FucT or XylT, or both, in a modified plant according to the invention is less than 95%, less than 90%, less than 80%, less than 70%, less than 60%, less than 50%, less than 40%, less than 30%, less than 20%, less than 10%, or less than 5% of the protein level of the same FucT or XylT in a plant that is not a mutant or that has not been genetically modified to inhibit the expression of that FucT or XylT, or both the FucT or XylT. The expression level of the FucT or XylT, or both, may be measured directly, for example, by assaying for the level of FucT or XylT, or both, expressed in the plant cell or plant, or indirectly, for example, by observing the effect in a transgenic plant at the phenotypic level, i.e., by transgenic plant analysis, observed as a reduction, or even elimination, of the attachment of  $\beta$ 1,2-xylose and/or  $\alpha$ 1,3-fucose residues to the glycoprotein *N*-glycans in the plant.

In other embodiments of the invention, the activity of FucT or XylT, or both, is reduced or eliminated by transforming a plant cell with an expression cassette comprising a polynucleotide encoding a polypeptide that inhibits the activity of FucT or XylT, or both. The activity of a FucT or XylT is inhibited according to the present invention if the activity of the FucT or XylT is statistically lower than the activity of the same FucT or XylT in a plant that has not been genetically modified to inhibit the activity of that FucT or XylT. In particular embodiments of the invention, the activity of the FucT or XylT in a modified plant according to the invention is less than 95%, less than 90%, less than 80%, less than 70%, less than 60%, less than 50%, less than



40%, less than 30%, less than 20%, less than 10%, or less than 5% of the activity of the same FucT or XylT in a plant that has not been genetically modified to inhibit the expression of that FucT or XylT. The activity of a FucT or XylT is "eliminated" according to the invention when it is not detectable by the assay methods described elsewhere herein.

In other embodiments, the activity of a FucT or XylT, or both, may be reduced or eliminated by disrupting the gene encoding the FucT or XylT, or both of these genes. The invention encompasses mutagenized plants, particularly plants that are members of the duckweed family, that carry mutations in a FucT or XylT gene, or mutations in both genes, where the mutations reduce expression of the FucT and/or XylT gene or inhibit the activity of the encoded FucT and/or XylT.

The methods of the invention can involve any method or mechanism known in the art for reducing or eliminating the activity or level of FucT and/or XylT in the cells of a higher plant, including, but not limited to, antisense suppression, sense suppression, RNA interference, directed deletion or mutation, dominant-negative strategies, and the like. Thus, the methods and compositions disclosed herein are not limited to any mechanism or theory of action and include any method where expression or function of FucT and/or XylT is inhibited in the cells of the higher plant of interest, thereby altering the *N*-glycosylation pattern of endogenous and heterologous glycoproteins produced in the plant.

For example, in some embodiments, the FucT inhibitory sequence or the XylT inhibitory sequence (or both) is expressed in the sense orientation, wherein the sense-oriented transcripts cause cosuppression of the expression of one or both of these enzymes. Alternatively, the FucT and/or XylT inhibitory sequence (e.g., full-length sequence, truncated sequence, fragments of the sequence, combinations thereof, and the like) can be expressed in the antisense orientation and thus inhibit endogenous FucT and/or XylT expression or function by antisense mechanisms.

In yet other embodiments, the FucT and/or XylT inhibitory sequence or sequences are expressed as a hairpin RNA, which comprises both a sense sequence and an antisense sequence. In embodiments comprising a hairpin structure, the loop structure may comprise any suitable nucleotide sequence including for example 5' untranslated and/or translated regions of the gene to be suppressed, such as the 5' UTR and/or translated region of the FucT polynucleotide of SEQ ID NO:1 or 2, or the



5' UTR and/or translated region of the XylT polynucleotide of SEQ ID NO:4, 5, 19, or 20, and the like. In some embodiments, the FucT or XylT inhibitory sequence expressed as a hairpin is encoded by an inverted region of the FucT or XylT nucleotide sequence. In yet other embodiments, the FucT and/or XylT inhibitory sequences are expressed as double-stranded RNA, where one FucT and/or XylT inhibitory sequence is expressed in the sense orientation and another complementary sequence is expressed in the antisense orientation. Double-stranded RNA, hairpin structures, and combinations thereof comprising FucT nucleotide sequences, XylT nucleotide sequences, or combinations thereof may operate by RNA interference, cosuppression, antisense mechanism, any combination thereof, or by means of any other mechanism that causes inhibition of FucT and/or XylT expression or function.

Thus, many methods may be used to reduce or eliminate the activity of a FucT or XylT, or both of these proteins, and any isoforms thereof. By "isoform" is intended a naturally occurring protein variant of the FucT or XylT protein of interest, where the variant is encoded by a different gene. Generally, isoforms of a particular FucT or XylT protein of interest are encoded by a nucleotide sequence having at least 90% sequence identity to the nucleotide sequence encoding the FucT or XylT protein of interest. More than one method may be used to reduce or eliminate the activity of a single plant FucT or XylT, and isoforms thereof. Non-limiting examples of methods of reducing or eliminating the activity of a plant FucT or XylT are given below.

*Polynucleotide-Based Methods:*

In some embodiments of the present invention, a plant cell is transformed with an expression cassette that is capable of expressing a polynucleotide that inhibits the expression of FucT or XylT, or both. The term "expression" as used herein refers to the biosynthesis of a gene product, including the transcription and/or translation of the gene product. For example, for the purposes of the present invention, an expression cassette capable of expressing a polynucleotide that inhibits the expression of at least one FucT or XylT, or both, is an expression cassette capable of producing an RNA molecule that inhibits the transcription and/or translation of at least one FucT or XylT, or both. The "expression" or "production" of a protein or polypeptide from a DNA molecule refers to the transcription and translation of the coding sequence to produce the protein or polypeptide, while the "expression" or "production" of a protein or

polypeptide from an RNA molecule refers to the translation of the RNA coding sequence to produce the protein or polypeptide.

Examples of polynucleotides that inhibit the expression of a FucT or XylT, or both, are given below.

5

Sense Suppression/Cosuppression

In some embodiments of the invention, inhibition of the expression of FucT or XylT, or both, may be obtained by sense suppression or cosuppression. For cosuppression, an expression cassette is designed to express an RNA molecule corresponding to all or part of a messenger RNA encoding a FucT or XylT, or both, in the "sense" orientation. Overexpression of the RNA molecule can result in reduced expression of the native gene. Accordingly, multiple plant lines transformed with the cosuppression expression cassette are screened to identify those that show the greatest inhibition of FucT or XylT expression.

15

The polynucleotide used for cosuppression may correspond to all or part of the sequence encoding the FucT or XylT, all or part of the 5' and/or 3' untranslated region of a FucT or XylT transcript, or all or part of both the coding sequence and the untranslated regions of a transcript encoding FucT or XylT. In some embodiments where the polynucleotide comprises all or part of the coding region for the FucT or XylT protein, the expression cassette is designed to eliminate the start codon of the polynucleotide so that no protein product will be transcribed.

20

Cosuppression may be used to inhibit the expression of plant genes to produce plants having undetectable protein levels for the proteins encoded by these genes. See, for example, Broin *et al.* (2002) *Plant Cell* 14:1417-1432. Cosuppression may also be used to inhibit the expression of multiple proteins in the same plant. See, for example, U.S. Patent No. 5,942,657. Methods for using cosuppression to inhibit the expression of endogenous genes in plants are described in Flavell *et al.* (1994) *Proc. Natl. Acad. Sci. USA* 91:3490-3496; Jorgensen *et al.* (1996) *Plant Mol. Biol.* 31:957-973; Johansen and Carrington (2001) *Plant Physiol.* 126:930-938; Broin *et al.* (2002) *Plant Cell* 14:1417-1432; Stoutjesdijk *et al.* (2002) *Plant Physiol.* 129:1723-1731; Yu *et al.* (2003) *Phytochemistry* 63:753-763; and U.S. Patent Nos. 5,034,323, 5,283,184, and 5,942,657; each of which is herein incorporated by reference. The efficiency of cosuppression may be increased by including a poly-dT region in the expression

30

cassette at a position 3' to the sense sequence and 5' of the polyadenylation signal. See, U.S. Patent Publication No. 20020048814, herein incorporated by reference. Typically, such a nucleotide sequence has substantial sequence identity to the sequence of the transcript of the endogenous gene, optimally greater than about 65% sequence identity, more optimally greater than about 85% sequence identity, most optimally greater than about 95% sequence identity. See, U.S. Patent Nos. 5,283,184 and 5,034,323; herein incorporated by reference.

#### Antisense Suppression

In some embodiments of the invention, inhibition of the expression of FucT or XylT, or both, may be obtained by antisense suppression. For antisense suppression, the expression cassette is designed to express an RNA molecule complementary to all or part of a messenger RNA encoding the FucT or XylT. Overexpression of the antisense RNA molecule can result in reduced expression of the native gene. Accordingly, multiple plant lines transformed with the antisense suppression expression cassette are screened to identify those that show the greatest inhibition of FucT or XylT expression.

The polynucleotide for use in antisense suppression may correspond to all or part of the complement of the sequence encoding the FucT or XylT, all or part of the complement of the 5' and/or 3' untranslated region of the FucT or XylT transcript, or all or part of the complement of both the coding sequence and the untranslated regions of a transcript encoding the FucT or XylT. In addition, the antisense polynucleotide may be fully complementary (i.e., 100% identical to the complement of the target sequence) or partially complementary (i.e., less than 100% identical to the complement of the target sequence) to the target sequence. Antisense suppression may be used to inhibit the expression of multiple proteins in the same plant. See, for example, U.S. Patent No. 5,942,657. Furthermore, portions of the antisense nucleotides may be used to disrupt the expression of the target gene. Generally, sequences of at least 50 nucleotides, 100 nucleotides, 200 nucleotides, 300, 400, 450, 500, 550, or greater may be used. Methods for using antisense suppression to inhibit the expression of endogenous genes in plants are described, for example, in Liu *et al.* (2002) *Plant Physiol.* 129:1732-1743 and U.S. Patent Nos. 5,759,829 and 5,942,657, each of which is herein incorporated by reference. Efficiency of antisense

suppression may be increased by including a poly-dT region in the expression cassette at a position 3' to the antisense sequence and 5' of the polyadenylation signal. See, U.S. Patent Publication No. 20020048814, herein incorporated by reference.

5                   Double-Stranded RNA Interference

In some embodiments of the invention, inhibition of the expression of a FucT or XylT, or both, may be obtained by double-stranded RNA (dsRNA) interference. For dsRNA interference, a sense RNA molecule like that described above for cosuppression and an antisense RNA molecule that is fully or partially  
10 complementary to the sense RNA molecule are expressed in the same cell, resulting in inhibition of the expression of the corresponding endogenous messenger RNA.

Expression of the sense and antisense molecules can be accomplished by designing the expression cassette to comprise both a sense sequence and an antisense sequence. Alternatively, separate expression cassettes may be used for the sense and  
15 antisense sequences. Multiple plant lines transformed with the dsRNA interference expression cassette or expression cassettes are then screened to identify plant lines that show the greatest inhibition of FucT or XylT expression. Methods for using dsRNA interference to inhibit the expression of endogenous plant genes are described in Waterhouse *et al.* (1998) *Proc. Natl. Acad. Sci. USA* 95:13959-13964, Liu *et al.*  
20 (2002) *Plant Physiol.* 129:1732-1743, and WO 99/49029, WO 99/53050, WO 99/61631, and WO 00/49035; each of which is herein incorporated by reference.

Hairpin RNA Interference and Intron-Containing Hairpin RNA Interference

In some embodiments of the invention, inhibition of the expression of FucT or  
25 XylT, or both, may be obtained by hairpin RNA (hpRNA) interference or intron-containing hairpin RNA (ihpRNA) interference. These methods are highly efficient at inhibiting the expression of endogenous genes. See, Waterhouse and Helliwell (2003) *Nat. Rev. Genet.* 4:29-38 and the references cited therein.

For hpRNA interference, the expression cassette is designed to express an  
30 RNA molecule that hybridizes with itself to form a hairpin structure that comprises a single-stranded loop region and a base-paired stem. The base-paired stem region comprises a sense sequence corresponding to all or part of the endogenous messenger RNA encoding the gene whose expression is to be inhibited, and an antisense

sequence that is fully or partially complementary to the sense sequence. Thus, the base-paired stem region of the molecule generally determines the specificity of the RNA interference. hpRNA molecules are highly efficient at inhibiting the expression of endogenous genes, and the RNA interference they induce is inherited by  
5 subsequent generations of plants. See, for example, Chuang and Meyerowitz (2000) *Proc. Natl. Acad. Sci. USA* 97:4985-4990; Stoutjesdijk *et al.* (2002) *Plant Physiol.* 129:1723-1731; and Waterhouse and Helliwell (2003) *Nat. Rev. Genet.* 4:29-38. Methods for using hpRNA interference to inhibit or silence the expression of genes are described, for example, in Chuang and Meyerowitz (2000) *Proc. Natl. Acad. Sci.*  
10 *USA* 97:4985-4990; Stoutjesdijk *et al.* (2002) *Plant Physiol.* 129:1723-1731; Waterhouse and Helliwell (2003) *Nat. Rev. Genet.* 4:29-38; Pandolfini *et al.* *BMC Biotechnology* 3:7, and U.S. Patent Publication No. 20030175965; each of which is herein incorporated by reference. A transient assay for the efficiency of hpRNA  
15 constructs to silence gene expression *in vivo* has been described by Panstruga *et al.* (2003) *Mol. Biol. Rep.* 30:135-140, herein incorporated by reference.

For ihpRNA, the interfering molecules have the same general structure as for hpRNA, but the RNA molecule additionally comprises an intron that is capable of being spliced in the cell in which the ihpRNA is expressed. The use of an intron minimizes the size of the loop in the hairpin RNA molecule following splicing, and  
20 this increases the efficiency of interference. See, for example, Smith *et al.* (2000) *Nature* 407:319-320. In fact, Smith *et al.* show 100% suppression of endogenous gene expression using ihpRNA-mediated interference. Methods for using ihpRNA interference to inhibit the expression of endogenous plant genes are described, for example, in Smith *et al.* (2000) *Nature* 407:319-320; Wesley *et al.* (2001) *Plant J.*  
25 27:581-590; Wang and Waterhouse (2001) *Curr. Opin. Plant Biol.* 5:146-150; Waterhouse and Helliwell (2003) *Nat. Rev. Genet.* 4:29-38; Helliwell and Waterhouse (2003) *Methods* 30:289-295, and U.S. Patent Publication No. 20030180945, each of which is herein incorporated by reference.

The expression cassette for hpRNA interference may also be designed such  
30 that the sense sequence and the antisense sequence do not correspond to an endogenous RNA. In this embodiment, the sense and antisense sequence flank a loop sequence that comprises a nucleotide sequence corresponding to all or part of the endogenous messenger RNA of the target gene. Thus, it is the loop region that



determines the specificity of the RNA interference. See, for example, WO 02/00904, herein incorporated by reference.

Transcriptional gene silencing (TGS) may be accomplished through use of hpRNA constructs wherein the inverted repeat of the hairpin shares sequence identity with the promoter region of a gene to be silenced. Processing of the hpRNA into short RNAs which can interact with the homologous promoter region may trigger degradation or methylation to result in silencing (Aufsatz *et al.* (2002) *PNAS* 99 (Suppl. 4):16499-16506; Mette *et al.* (2000) *EMBO J* 19(19):5194-5201).

Expression cassettes that are designed to express an RNA molecule that forms a hairpin structure are referred to herein as RNAi expression cassettes. In some embodiments, the RNAi expression cassette is designed in accordance with a strategy outlined in Figure 28. In such embodiments, an RNAi expression cassette can be designed to suppress the expression of the individual FucT and XylT genes (i.e., each cassette provides a single gene knockout), or can be designed to suppress the expression of both the FucT and XylT genes (i.e., a single RNAi expression cassette expresses an inhibitory molecule that provides for suppression of expression of both of these genes). Where the RNAi expression cassette suppresses expression of both the FucT and XylT genes, it is referred to herein as a "chimeric" RNAi expression cassette. The single-gene and chimeric RNAi expression cassettes can be designed to express larger hpRNA structures or, alternatively, small hpRNA structures, as noted herein below.

Thus, in some embodiments, the RNAi expression cassette is designed to express larger hpRNA structures having sufficient homology to the target mRNA transcript to provide for post-transcriptional gene silencing of one or both of a FucT and XylT gene. For larger hp RNA structures, the sense strand of the RNAi expression cassette is designed to comprise in the 5'-to-3' direction the following operably linked elements: a promoter of interest, a forward fragment of the FucT or XylT gene sequence comprising about 500 to about 800 nucleotides (nt) of a sense strand for FucT or XylT, respectively, a spacer sequence comprising about 100 to about 700 nt of any sequence as noted herein below, and a reverse fragment of the XylT or FucT gene sequence, wherein the reverse fragment comprises the antisense sequence complementary to the respective (i.e., FucT or XylT) forward fragment. Thus, for example, if a forward fragment is represented by nucleotides "...acttg...",



the corresponding reverse fragment is represented by nucleotides "...caagt...", and the sense strand for such an RNAi expression cassette would comprise the following sequence: "5'...acttg...nnnn...caagt...-3'", where "nnnn" represents the spacer sequence.

5           It is recognized that the forward fragment can comprise a nucleotide sequence that is 100% identical to the corresponding portion of the sense strand for the target FucT or XylT gene sequence, or in the alternative, can comprise a sequence that shares at least 90%, at least 95%, or at least 98% sequence identity to the corresponding portion of the sense strand for the target FucT or XylT gene to be  
10           silenced. In like manner, it is recognized that the reverse fragment does not have to share 100% sequence identity to the complement of the forward fragment; rather it must be of sufficient length and sufficient complementarity to the forward fragment sequence such that when the inhibitory RNA molecule is expressed, the transcribed regions corresponding to the forward fragment and reverse fragment will hybridize to  
15           form the base-paired stem (i.e., double-stranded portion) of the hp RNA structure. By "sufficient length" is intended a length that is at least 10%, at least 15%, at least 20%, at least 30%, at least 40% of the length of the forward fragment, more frequently at least 50%, at least 75%, at least 90%, or least 95% of the length of the forward fragment. By "sufficient complementarity" is intended the sequence of the reverse  
20           fragment shares at least 90%, at least 95%, at least 98% sequence identity with the complement of that portion of the forward fragment that will hybridize with the reverse fragment to form the base-paired stem of the hp RNA structure. Thus, in some embodiments, the reverse fragment is the complement (i.e., antisense version) of the forward fragment.

25           In designing such an RNAi expression cassette, the lengths of the forward fragment, spacer sequence, and reverse fragments are chosen such that the combined length of the polynucleotide that encodes the hpRNA construct is about 650 to about 2500 nt, about 750 to about 2500 nt, about 750 to about 2400 nt, about 1000 to about 2400 nt, about 1200 to about 2300 nt, about 1250 to about 2100 nt, or about 1500 to  
30           about 1800. In some embodiments, the combined length of the expressed hairpin construct is about 650 nt, about 700 nt, about 750 nt, about 800 nt, about 850 nt, about 900 nt, about 950 nt, about 1000 nt, about 1050 nt, about 1100 nt, about 1150 nt, about 1200 nt, about 1250 nt, about 1300 nt, about 1350 nt, about 1400 nt, about 1450

nt, about 1500 nt, about 1550 nt, about 1600 nt, about 1650 nt, about 1700 nt, about 1750 nt, about 1800 nt, about 1850 nt, about 1900 nt, about 1950 nt, about 2000 nt, about 2050 nt, about 2100 nt, about 2150 nt, about 2200 nt, about 2250 nt, about 2300 nt, or any such length between about 650 nt to about 2300 nt.

5           In some embodiments, the forward fragment comprises about 500 to about 800 nt, for example, 500, 525, 550, 575, 600, 625, 650, 675, 700, 725, 750, 775, or 800 nt of a sense strand for FucT or XylT, for example, of the sense strand set forth in SEQ ID NO:1 or 2 (FucT) or SEQ ID NO:4, 5, 19, or 20 (XylT); the spacer sequence comprises about 100 to about 700 nt, for example, 100, 125, 150, 175, 200, 225, 250,  
10       275, 300, 325, 350, 375, 400, 425, 450, 475, 500, 525, 550, 575, 600, 625, 650, 675, or 700 nt of any sequence as noted below, and the reverse fragment comprises the antisense strand for the forward fragment sequence, or a sequence having sufficient length and sufficient complementarity to the forward fragment sequence.

          The spacer sequence can be any sequence that has insufficient homology to  
15       the target gene, i.e., FucT or XylT, and insufficient homology to itself such that the portion of the expressed inhibitory RNA molecule corresponding to the spacer region fails to self-hybridize, and thus forms the loop of the hairpin RNA structure. In some embodiments, the spacer sequence comprises an intron, and thus the expressed inhibitory RNA molecule forms an ihpRNA as noted herein above. In other  
20       embodiments, the spacer sequence comprises a portion of the sense strand for the FucT or XylT gene to be silenced, for example, a portion of the sense strand set forth in SEQ ID NO:1 or 2 (FucT) or SEQ ID NO:4, 5, 19, or 20 (XylT), particularly a portion of the sense strand immediately downstream from the forward fragment sequence.

25           The operably linked promoter can be any promoter of interest that provides for expression of the operably linked inhibitory polynucleotide within the plant of interest, including one of the promoters disclosed herein below. The regulatory region can comprise additional regulatory elements that enhance expression of the inhibitory polynucleotide, including, but not limited to, the 5' leader sequences and 5' leader sequences plus plant introns discussed herein below.  
30

          In one embodiment, the RNAi expression cassette is designed to suppress expression of the FucT polypeptide of SEQ ID NO:3, a biologically active variant of the FucT polypeptide of SEQ ID NO:3, or a FucT polypeptide encoded by a sequence

having at least 75% sequence identity to the sequence of SEQ ID NO:1 or SEQ ID NO:2, for example, at least 75%, at least 80%, at least 85%, at least 90%, or at least 95% sequence identity to the sequence of SEQ ID NO:1 or SEQ ID NO:2. In this manner, the sense strand of the RNAi expression cassette is designed to comprise in the 5'-to-3' direction the following operably linked elements: a promoter of interest; a forward fragment of the FucT gene sequence, wherein the forward fragment comprises nt 255-985 of SEQ ID NO:1; a spacer sequence comprising about 100 to about 700 nt of any sequence as noted above; and a reverse fragment of the FucT gene sequence, wherein the reverse fragment comprises the complement (i.e., antisense version) of nt 255-985 of SEQ ID NO:1. In one such embodiment, the spacer sequence is represented by nt 986-1444 of SEQ ID NO:1, and the total length of that portion of the sense strand of the RNAi expression cassette corresponding to the coding sequence for the hpRNA structure is 1918 nt. Stably transforming a plant with a nucleotide construct comprising this RNAi expression cassette, for example, the vector shown in Figure 8, effectively inhibits expression of FucT within the plant cells of the plant in which the hpRNA structure is expressed. In one embodiment, the plant of interest is a member of the duckweed family, for example, a member of the *Lemnaceae*, and the plant has been stably transformed with the vector shown in Figure 8.

In other embodiments of the invention, the RNAi expression cassette is designed to suppress expression of the XylT polypeptide of SEQ ID NO:6 or SEQ ID NO:21, a biologically active variant of the XylT polypeptide of SEQ ID NO:6 or SEQ ID NO:21, or a XylT polypeptide encoded by a sequence having at least 75% sequence identity to the sequence of SEQ ID NO:4, SEQ ID NO:5, SEQ ID NO:19, or SEQ ID NO:20, for example, at least 75%, at least 80%, at least 85%, at least 90%, or at least 95% sequence identity to the sequence of SEQ ID NO:4, SEQ ID NO:5, SEQ ID NO:19, or SEQ ID NO:20. In this manner, the sense strand of the RNAi expression cassette is designed to comprise in the 5'-to-3' direction the following operably linked elements: a promoter of interest; a forward fragment of the XylT gene sequence, wherein the forward fragment comprises nt 318-1052 of SEQ ID NO:4; a spacer sequence comprising about 100 to about 700 nt of any sequence as noted above; and a reverse fragment of the XylT gene sequence, wherein the reverse fragment comprises the complement (i.e., antisense version) of nt 318-1052 of SEQ

ID NO:4. In one such embodiment, the spacer sequence is represented by nt 1053-1599 of SEQ ID NO:4, and the total length of that portion of the sense strand of the RNAi expression cassette corresponding to the coding sequence for the hpRNA structure is 2015 nt. Stably transforming a plant with a nucleotide construct comprising this RNAi expression cassette, for example, the vector shown in Figure 9, effectively inhibits expression of FucT within the plant cells of the plant in which the hpRNA structure is expressed. In one embodiment, the plant of interest is a member of the duckweed family, for example, a member of the *Lemnaceae*, and the plant has been stably transformed with the vector shown in Figure 9.

In other embodiments, the sense strand of the RNAi expression cassette is designed to comprise in the 5'-to-3' direction the following operably linked elements: a promoter of interest; a forward fragment of the XylT gene sequence, wherein the forward fragment comprises nt 1-734 of SEQ ID NO:19; a spacer sequence comprising about 100 to about 700 nt of any sequence as noted above; and a reverse fragment of the XylT gene sequence, wherein the reverse fragment comprises the complement (i.e., antisense version) of nt 1-734 of SEQ ID NO:19. In one such embodiment, the spacer sequence is represented by nt 735-1282 of SEQ ID NO:19, and the total length of that portion of the sense strand of the RNAi expression cassette corresponding to the coding sequence for the hpRNA structure is 2015 nt. Stably transforming a plant with a nucleotide construct comprising this RNAi expression cassette, for example, the vector shown in Figure 9, effectively inhibits expression of FucT within the plant cells of the plant in which the hpRNA structure is expressed. In one embodiment, the plant of interest is a member of the duckweed family, for example, a member of the *Lemnaceae*, and the plant has been stably transformed with the vector shown in Figure 9.

In yet other embodiments, larger hpRNA structures can be designed such that the antisense and sense sequences are in opposite orientation. In this manner, the sense strand of the RNAi expression cassette is designed to comprise in the 5'-to-3' direction the following operably linked elements: a promoter of interest, the full-length FucT or XylT gene sequence in the antisense orientation, and a forward fragment of the FucT or XylT gene sequence comprising the 3'-half of the sequence in the sense orientation (see Figure 28, design 1). In this type of construct, the 3'-half of the sequence forms the base-paired (i.e., double-stranded) stem of the hpRNA, and the

5'-half of the sequence acts as a spacer sequence. Without being bound by any theory or mechanism, the 3' region of the FucT or XylT sequence is chosen to form the double-stranded region of the hpRNA for for this construct because this region is relatively conserved among different plant species compared to the 5' region.

5           In one such embodiment, the RNAi expression cassette is designed to suppress expression of the FucT polypeptide of SEQ ID NO:3, a biologically active variant of the FucT polypeptide of SEQ ID NO:3, or a FucT polypeptide encoded by a sequence having at least 75% sequence identity to the sequence of SEQ ID NO:1 or SEQ ID NO:2, for example, at least 75%, at least 80%, at least 85%, at least 90%, or at least 10   95% sequence identity to the sequence of SEQ ID NO:1 or SEQ ID NO:2. In this manner, the sense strand of the RNAi expression cassette is designed to comprise in the 5'-to-3' direction the following operably linked elements: a promoter of interest; nucleotides 1-1865 of SEQ ID NO:1 in antisense orientation, and nucleotides 950-1865 of SEQ ID NO:1 in the sense orientation. Stably transforming a plant with a 15   nucleotide construct comprising this RNAi expression cassette effectively inhibits expression of FucT within the plant cells of the plant in which the hpRNA structure is expressed. In one embodiment, the plant of interest is a member of the duckweed family, for example, a member of the *Lemnaceae*.

          In another such embodiment, the RNAi expression cassette is designed to 20   suppress expression of the XylT polypeptide of SEQ ID NO:6 or SEQ ID NO:21, a biologically active variant of the XylT polypeptide of SEQ ID NO:6 or SEQ ID NO:21, or a XylT polypeptide encoded by a sequence having at least 75% sequence identity to the sequence of SEQ ID NO:4, SEQ ID NO:5, SEQ ID NO:19, or SEQ ID NO:20, for example, at least 75%, at least 80%, at least 85%, at least 90%, or at least 25   95% sequence identity to the sequence of SEQ ID NO:5, SEQ ID NO:19, or SEQ ID NO:20. In this manner, the sense strand of the RNAi expression cassette is designed to comprise in the 5'-to-3' direction the following operably linked elements: a promoter of interest, nucleotides 1-1860 of SEQ ID NO:4 in antisense orientation, and nucleotides 950-1860 of SEQ ID NO:4 in the sense orientation. Stably transforming a 30   plant with a nucleotide construct comprising this RNAi expression cassette effectively inhibits expression of XylT within the plant cells of the plant in which the hpRNA structure is expressed. In one embodiment, the plant of interest is a member of the duckweed family, for example, a member of the *Lemnaceae*.



In another such embodiment, the sense strand of the RNAi expression cassette is designed to comprise in the 5'-to-3' direction the following operably linked elements: a promoter of interest, nucleotides 1-1282 of SEQ ID NO:19 in antisense orientation, and nucleotides 652-1282 of SEQ ID NO:19 in the sense orientation.

5 Stably transforming a plant with a nucleotide construct comprising this RNAi expression cassette effectively inhibits expression of XylT within the plant cells of the plant in which the hpRNA structure is expressed. In one embodiment, the plant of interest is a member of the duckweed family, for example, a member of the *Lemnaceae*.

10 Where suppression of both the FucT and XylT proteins is desired, it can be achieved by introducing these single-gene RNAi expression cassettes into the plant in a single transformation event, for example, by assembling these single-gene RNAi expression cassettes within a single transformation vector, for example, a vector similar to that shown in Figure 11, or as separate co-transformation events, for  
15 example, by assembling these single-gene RNAi expression cassettes within two transformation vectors, for example, vectors similar to those shown in Figures 8 and 9, using any suitable transformation method known in the art, including but not limited to the transformation methods disclosed elsewhere herein.

Alternatively, suppression of both the FucT and XylT proteins can be achieved  
20 by introducing into the higher plant of interest a chimeric RNAi expression cassette as noted herein above. Thus, in some embodiments of the invention, the sense strand of a chimeric RNAi expression cassette is designed to comprise in the 5'-to-3' direction the following operably linked elements: a promoter of interest; a chimeric forward fragment, comprising about 500 to about 650 nucleotides (nt) of a sense strand for  
25 FucT and about 500 to about 650 nt of a sense strand for XylT, wherein the FucT sequence and XylT sequence can be in either order; a spacer sequence comprising about 100 to about 700 nt of any sequence; and a reverse fragment of the chimeric forward fragment, wherein the reverse fragment comprises the antisense sequence complementary to the respective chimeric forward fragment.

30 As previously noted for the individual RNAi expression cassettes, it is recognized that the individual FucT or XylT sequence within the chimeric forward fragment can comprise a nucleotide sequence that is 100% identical to the corresponding portion of the sense strand for the target FucT and XylT gene



sequence, respectively, or in the alternative, can comprise a sequence that shares at least 90%, at least 95%, or at least 98% sequence identity to the corresponding portion of the sense strand for the target FucT or XylT gene to be silenced. In like manner, it is recognized that the reverse fragment does not have to share 100% sequence identity to the complement of the chimeric forward fragment; rather it must be of sufficient length and sufficient complementarity to the chimeric forward fragment sequence, as defined herein above, such that when the inhibitory RNA molecule is expressed, the transcribed regions corresponding to the chimeric forward fragment and reverse fragment will hybridize to form the base-paired stem (i.e., double-stranded portion) of the hpRNA structure. In designing such a chimeric RNAi expression cassette, the lengths of the forward fragment, spacer sequence, and reverse fragments are chosen such that the combined length of the polynucleotide that encodes the hpRNA structure is about 1200 to about 3300 nt, about 1250 to about 3300 nt, about 1300 to about 3300 nt, about 1350 to about 3300 nt, about 1400 to about 3300 nt, about 1450 nt to about 3300 nt, about 1500 to about 3300 nt, about 2200 to about 3100 nt, about 2250 to about 2800 nt, or about 2500 to about 2700 nt. In some embodiments, the combined length of the expressed hairpin construct is about 1200 nt, about 1250 nt, about 1300 nt, about 1350 nt, about 1400 nt, about 1450 nt, about 1500 nt, about 1800 nt, about 2200 nt, about 2250 nt, about 2300 nt, about 2350 nt, about 2400 nt, about 2450 nt, about 2500 nt, about 2550 nt, about 2600 nt, about 2650 nt, about 2700 nt, about 2750 nt, about 2800 nt, about 2850 nt, about 2900 nt, about 2950 nt, about 3000 nt, about 3050 nt, about 3100 nt, about 3150 nt, about 3200 nt, about 3250 nt, about 3300 nt, or any such length between about 1200 nt to about 3300 nt.

In some embodiments, the chimeric forward fragment comprises about 500 to about 650 nt, for example, 500, 525, 550, 575, 600, 625, or 650 nt, of a sense strand for FucT, for example, of the sense strand set forth in SEQ ID NO:1 or 2, and about 500 to about 650 nt, for example, 500, 525, 550, 575, 600, 625, or 650 nt, of a sense strand for XylT, for example, of the sense strand set forth in SEQ ID NO:4, 5, 19, or 20, where the FucT and XylT sequence can be fused in either order; the spacer sequence comprises about 100 to about 700 nt, for example, 100, 200, 225, 250, 275, 300, 325, 350, 375, 400, 425, 450, 475, 500, 525, 550, 575, 600, 625, 650, 675, or 700 nt of any sequence of interest; and the reverse fragment comprises the antisense strand

for the chimeric forward fragment sequence, or a sequence having sufficient length and sufficient complementarity to the chimeric forward fragment sequence.

As noted above for the single-gene RNAi expression cassettes, the spacer sequence can be any sequence that has insufficient homology to the target gene, i.e.,  
5 FucT or XylT, and insufficient homology to itself such that the portion of the expressed inhibitory RNA molecule corresponding to the spacer region fails to self-hybridize, and thus forms the loop of the hpRNA structure. In some embodiments, the spacer sequence comprises an intron, and thus the expressed inhibitory RNA molecule forms an ihpRNA as noted herein above. In other embodiments, the spacer  
10 sequence comprises a portion of the sense strand for the FucT or XylT gene to be silenced, for example, a portion of the sense strand set forth in SEQ ID NO:1 or 2 (FucT) or SEQ ID NO:4, 5, 19, or 20 (XylT). In one embodiment, the chimeric forward fragment comprises the FucT and XylT sequence fused in that order, and the spacer sequence comprises a portion of the XylT sense strand immediately  
15 downstream from the XylT sequence contained within the chimeric forward fragment. In another embodiment, the chimeric forward fragment comprises the XylT and FucT sequence fused in that order, and the spacer sequence comprises a portion of the FucT sense strand immediately downstream from the FucT sequence contained within the chimeric forward fragment.

20 In some embodiments, the chimeric RNAi expression cassette is designed to suppress expression of the FucT polypeptide of SEQ ID NO:3, a biologically active variant of the FucT polypeptide of SEQ ID NO:3, or a FucT polypeptide encoded by a sequence having at least 75% sequence identity to the sequence of SEQ ID NO:1 or  
25 SEQ ID NO:2, for example, at least 75%, at least 80%, at least 85%, at least 90%, or at least 95% sequence identity to the sequence of SEQ ID NO:1 or SEQ ID NO:2, and to suppress expression of the XylT polypeptide of SEQ ID NO:6 or SEQ ID NO:21, a biologically active variant of the XylT polypeptide of SEQ ID NO:6 or SEQ ID  
30 NO:21, or a XylT polypeptide encoded by a sequence having at least 75% sequence identity to the sequence of SEQ ID NO:4, SEQ ID NO:5, SEQ ID NO:19, or SEQ ID NO:20, for example, at least 75%, at least 80%, at least 85%, at least 90%, or at least 95% sequence identity to the sequence of SEQ ID NO:4, SEQ ID NO:5, SEQ ID NO:19, or SEQ ID NO:20. For some of these embodiments, the FucT sequence within the chimeric forward fragment is chosen such that it corresponds to nt 700 to nt

1400 of SEQ ID NO:1 or SEQ ID NO:2, and/or the XylT sequence within the chimeric forward fragment is chosen such that it corresponds to nt 700 to nt 1400 of SEQ ID NO:4 or SEQ ID NO:5, or nt 383 to 1083 of SEQ ID NO:19 or 20. Without being bound by theory, it is believed that this region (particularly for FucT) is  
5 relatively conserved among different plant species, and therefore is a potentially good target.

In other embodiments, the sense strand of the chimeric RNAi expression cassette is designed to comprise in the 5'-to-3' direction the following operably linked elements: a promoter of interest; a chimeric forward fragment comprising nt 254-855  
10 of SEQ ID NO:1 (FucT sequence) and nt 318-943 of SEQ ID NO:4 (XlyT sequence); a spacer sequence comprising about 100 to about 700 nt of any sequence as noted above; and a reverse fragment comprising the complement (i.e., antisense version) of the chimeric forward fragment, i.e., comprising the complement of nt 318-943 of SEQ ID NO:4 and the complement of nt 254-855 of SEQ ID NO:1. In a particular  
15 embodiment, the spacer sequence within this chimeric RNAi expression cassette is represented by nt 944-1443 of SEQ ID NO:4, and the total length of that portion of the sense strand of the RNAi expression cassette corresponding to the coding sequence for the hpRNA structure is 2956 nt.

In another such embodiment, the sense strand of the chimeric RNAi  
20 expression cassette is designed to comprise in the 5'-to-3' direction the following operably linked elements: a promoter of interest; a chimeric forward fragment comprising nt 318-943 of SEQ ID NO:4 (XlyT sequence) and nt 254-855 of SEQ ID NO:1 (FucT sequence); a spacer sequence comprising about 100 to about 700 nt of any sequence as noted above; and a reverse fragment comprising the complement  
25 (i.e., antisense version) of the chimeric forward fragment, i.e., comprising the complement of nt 254-855 of SEQ ID NO:1 and the complement of nt 318-943 of SEQ ID NO:4. In a particular embodiment, the spacer sequence within this chimeric RNAi expression cassette is represented by nt 856-1355 of SEQ ID NO:1, and the total length of that portion of the sense strand of the RNAi expression cassette  
30 corresponding to the coding sequence for the hpRNA structure is 2956 nt.

In yet other embodiments, the sense strand of the chimeric RNAi expression cassette is designed to comprise in the 5'-to-3' direction the following operably linked elements: a promoter of interest; a chimeric forward fragment comprising nt 254-855

of SEQ ID NO:1 (FucT sequence) and nt 1-626 of SEQ ID NO:19 (XlyT sequence); a spacer sequence comprising about 100 to about 700 nt of any sequence as noted above; and a reverse fragment comprising the complement (i.e., antisense version) of the chimeric forward fragment, i.e., comprising the complement of nt 1-626 of SEQ ID NO:19 and the complement of nt 254-855 of SEQ ID NO:1. In a particular embodiment, the spacer sequence within this chimeric RNAi expression cassette is represented by nt 627-1126 of SEQ ID NO:19, and the total length of that portion of the sense strand of the RNAi expression cassette corresponding to the coding sequence for the hpRNA structure is 2956 nt.

In another such embodiment, the sense strand of the chimeric RNAi expression cassette is designed to comprise in the 5'-to-3' direction the following operably linked elements: a promoter of interest; a chimeric forward fragment comprising nt 1-626 of SEQ ID NO:19 (XlyT sequence) and nt 254-855 of SEQ ID NO:1 (FucT sequence); a spacer sequence comprising about 100 to about 700 nt of any sequence as noted above; and a reverse fragment comprising the complement (i.e., antisense version) of the chimeric forward fragment, i.e., comprising the complement of nt 254-855 of SEQ ID NO:1 and the complement of nt 1-626 of SEQ ID NO:19. In a particular embodiment, the spacer sequence within this chimeric RNAi expression cassette is represented by nt 856-1355 of SEQ ID NO:1, and the total length of that portion of the sense strand of the RNAi expression cassette corresponding to the coding sequence for the hpRNA structure is 2956 nt.

Stably transforming a plant with a nucleotide construct comprising a chimeric RNAi expression cassette described herein, for example, stable transformation with the vector shown in Figure 10, effectively inhibits expression of both FucT and XylT within the plant cells of the plant in which the hpRNA structure is expressed. In one embodiment, the plant of interest is a member of the duckweed family, for example, a member of the *Lemnaceae*, and the plant has been stably transformed with the vector shown in Figure 10.

It is recognized that the plant can be stably transformed with at least two of these chimeric RNAi expression cassettes to provide for very efficient gene silencing of the FucT and XylT proteins, including silencing of any isoforms of these two proteins. See, for example, the two orientations provided in "possible design 2" of Figure 28. In this manner, the plant can be stably transformed with a first chimeric

RNAi expression cassette wherein the chimeric forward fragment comprises the FucT and XylT sequence fused in that order, and the spacer sequence comprises a portion of the XylT sense strand immediately downstream from the XylT sequence contained within the chimeric forward fragment; and with a second chimeric RNAi expression  
5 cassette wherein the chimeric forward fragment comprises the XylT and FucT sequence fused in that order, and the spacer sequence comprises a portion of the FucT sense strand immediately downstream from the FucT sequence contained within the chimeric forward fragment.

The operably linked promoter within any of the RNAi expression cassettes  
10 encoding large hpRNA structures, or large ihpRNA structures can be any promoter of interest that provides for expression of the operably linked inhibitory polynucleotide within the plant of interest, including one of the promoters disclosed herein below. The regulatory region can comprise additional regulatory elements that enhance expression of the inhibitory polynucleotide, including, but not limited to, the 5' leader  
15 sequences and 5' leader sequences plus plant introns discussed herein below.

In yet other embodiments, the RNAi expression cassette can be designed to provide for expression of small hpRNA structures having a base-paired stem region comprising about 200 base pairs or less. Expression of the small hpRNA structure is preferably driven by a promoter recognized by DNA-dependent RNA polymerase III.  
20 See, for example, U.S. Patent Application No. 20040231016, herein incorporated by reference in its entirety.

In this manner, the RNAi expression cassette is designed such that the transcribed DNA region encodes an RNA molecule comprising a sense and antisense nucleotide region, where the sense nucleotide sequence comprises about 19  
25 contiguous nucleotides having about 90% to about 100% sequence identity to a nucleotide sequence of about 19 contiguous nucleotides from the RNA transcribed from the gene of interest and where the antisense nucleotide sequence comprises about 19 contiguous nucleotides having about 90% to about 100% sequence identity to the complement of a nucleotide sequence of about 19 contiguous nucleotides of the  
30 sense sequence. The sense and antisense nucleotide sequences of the RNA molecule should be capable of forming a base-paired (i.e., double-stranded) stem region of RNA of about 19 to about 200 nucleotides, alternatively about 21 to about 90 or 100 nucleotides, or alternatively about 40 to about 50 nucleotides in length. However, the



length of the base-paired stem region of the RNA molecule may also be about 30, about 60, about 70 or about 80 nucleotides in length. Where the base-paired stem region of the RNA molecule is larger than 19 nucleotides, there is only a requirement that there is at least one double-stranded region of about 19 nucleotides (wherein there can be about one mismatch between the sense and antisense region) the sense strand of which is "identical" (allowing for one mismatch) with 19 consecutive nucleotides of the target FucT or XylT polynucleotide of interest. The transcribed DNA region of this type of RNAi expression cassette may comprise a spacer sequence positioned between the sense and antisense encoding nucleotide region. The spacer sequence is not related to the targeted FucT or XylT polynucleotide, and can range in length from 3 to about 100 nucleotides or alternatively from about 6 to about 40 nucleotides. This type of RNAi expression cassette also comprises a terminator sequence recognized by the RNA polymerase III, the sequence being an oligo dT stretch, positioned downstream from the antisense-encoding nucleotide region of the cassette. By "oligo dT stretch" is a stretch of consecutive T-residues. It should comprise at least 4 T-residues, but obviously may contain more T-residues.

It is recognized that in designing the short hpRNA, the fragments of the targeted gene sequence (i.e., fragments of FucT or XylT gene sequence) and any spacer sequence to be included within the hpRNA-encoding portion of the RNAi expression cassette are chosen to avoid GC-rich sequences, particularly those with three consecutive G/C's, and to avoid the occurrence of four or more consecutive T's or A's, as the string "TTTT..." serves as a terminator sequence recognized by the RNA polymerase III.

Thus, where gene silencing with a short hpRNA is desired, the RNAi expression cassette can be designed to comprise in the 5'-to-3' direction the following operably linked elements: a promoter recognized by a DNA dependent RNA polymerase III of the plant cell, as defined herein below; a DNA fragment comprising a sense and antisense nucleotide sequence, wherein the sense nucleotide sequence comprises at least 19 contiguous nucleotides having about 90% to about 100% sequence identity to a nucleotide sequence of at least 19 contiguous nucleotides from the sense strand of the FucT or XylT gene of interest, and wherein the antisense nucleotide sequence comprises at least 19 contiguous nucleotides having about 90% to about 100% sequence identity to the complement of a nucleotide sequence of at



least 19 contiguous nucleotides of the sense sequence, wherein the sense and antisense nucleotide sequence are capable of forming a double-stranded RNA of about 19 to about 200 nucleotides in length; and an oligo dT stretch comprising at least 4 consecutive T-residues.

5           In some embodiments of the invention, the RNAi expression cassette is designed to express a small hpRNA that suppresses expression of the FucT polypeptide of SEQ ID NO:3, a biologically active variant of the FucT polypeptide of SEQ ID NO:3, or a FucT polypeptide encoded by a sequence having at least 90% sequence identity to the sequence of SEQ ID NO:1 or SEQ ID NO:2. In this manner,  
10           the RNAi expression cassette can be designed to comprise in the 5'-to-3' direction the following operably linked elements: a promoter recognized by a DNA dependent RNA polymerase III of the plant cell, as defined herein below; a DNA fragment comprising a sense and antisense nucleotide sequence, wherein the sense nucleotide sequence comprises at least 19 contiguous nucleotides having about 90% to about  
15           100% sequence identity to a nucleotide sequence of at least 19 contiguous nucleotides of SEQ ID NO:1, and wherein the antisense nucleotide sequence comprises at least 19 contiguous nucleotides having about 90% to about 100% sequence identity to the complement of a nucleotide sequence of at least 19 contiguous nucleotides of the sense sequence, wherein the sense and antisense nucleotide sequence are capable of  
20           forming a double-stranded RNA of about 19 to about 200 nucleotides in length; and an oligo dT stretch comprising at least 4 consecutive T-residues.

          In other embodiments of the invention, the RNAi expression cassette is designed to express a small hpRNA that suppresses expression of the XylT polypeptide of SEQ ID NO:6 or SEQ ID NO:21, a biologically active variant of the  
25           XylT polypeptide of SEQ ID NO:6 or SEQ ID NO:21, or a XylT polypeptide encoded by a sequence having at least 90% sequence identity to the sequence of SEQ ID NO:4, SEQ ID NO:5, SEQ ID NO:19, or SEQ ID NO:20. In this manner, the RNAi expression cassette can be designed to comprise in the 5'-to-3' direction the following operably linked elements: a promoter recognized by a DNA dependent RNA  
30           polymerase III of the plant cell, as defined herein below; a DNA fragment comprising a sense and antisense nucleotide sequence, wherein the sense nucleotide sequence comprises at least 19 contiguous nucleotides having about 90% to about 100% sequence identity to a nucleotide sequence of at least 19 contiguous nucleotides of

SEQ ID NO:4, and wherein the antisense nucleotide sequence comprises at least 19 contiguous nucleotides having about 90% to about 100% sequence identity to the complement of a nucleotide sequence of at least 19 contiguous nucleotides of the sense sequence, wherein the sense and antisense nucleotide sequence are capable of forming a double stranded RNA of about 19 to about 200 nucleotides in length; and an oligo dT stretch comprising at least 4 consecutive T-residues.

#### Amplicon-Mediated Interference

Amplicon expression cassettes comprise a plant virus-derived sequence that contains all or part of the target gene but generally not all of the genes of the native virus. The viral sequences present in the transcription product of the expression cassette allow the transcription product to direct its own replication. The transcripts produced by the amplicon may be either sense or antisense relative to the target sequence (i.e., the messenger RNA for FucT or XylT, or both). Methods of using amplicons to inhibit the expression of endogenous plant genes are described, for example, in Angell and Baulcombe (1997) *EMBO J.* 16:3675-3684, Angell and Baulcombe (1999) *Plant J.* 20:357-362, and U.S. Patent No. 6,646,805, each of which is herein incorporated by reference.

#### Ribozymes

In some embodiments, the polynucleotide expressed by the expression cassette of the invention is catalytic RNA or has ribozyme activity specific for the messenger RNA of FucT or XylT, or both. Thus, the polynucleotide causes the degradation of the endogenous messenger RNA, resulting in reduced expression of the FucT or XylT, or both. This method is described, for example, in U.S. Patent No. 4,987,071, herein incorporated by reference.

#### Small Interfering RNA or Micro RNA

In some embodiments of the invention, inhibition of the expression of FucT or XylT, or both, may be obtained by RNA interference by expression of a gene encoding a micro RNA (miRNA). miRNAs are regulatory agents consisting of about 22 ribonucleotides. miRNA are highly efficient at inhibiting the expression of

endogenous genes. See, for example Javier *et al.* (2003) *Nature* 425: 257-263, herein incorporated by reference.

For miRNA interference, the expression cassette is designed to express an RNA molecule that is modeled on an endogenous miRNA gene. The miRNA gene  
5 encodes an RNA that forms a hairpin structure containing a 22-nucleotide sequence that is complementary to another endogenous gene (target sequence). For suppression of FucT or XylT expression, the 22-nucleotide sequence is selected from a FucT or XylT transcript sequence and contains 22 nucleotides of said FucT or XylT sequence in sense orientation and 21 nucleotides of a corresponding antisense sequence that is  
10 complementary to the sense sequence. miRNA molecules are highly efficient at inhibiting the expression of endogenous genes, and the RNA interference they induce is inherited by subsequent generations of plants.

#### Polypeptide-Based Inhibition of Gene Expression

15 In one embodiment, the polynucleotide encodes a zinc finger protein that binds to a gene encoding a FucT or XylT, or both, resulting in reduced expression of the gene. In particular embodiments, the zinc finger protein binds to a regulatory region of a FucT or XylT gene. In other embodiments, the zinc finger protein binds to a messenger RNA encoding a FucT or XylT and prevents its translation. Methods of  
20 selecting sites for targeting by zinc finger proteins have been described, for example, in U.S. Patent No. 6,453,242, and methods for using zinc finger proteins to inhibit the expression of genes in plants are described, for example, in U.S. Patent Publication No. 20030037355; each of which is herein incorporated by reference.

#### Polypeptide-Based Inhibition of Protein Activity

25 In some embodiments of the invention, the polynucleotide encodes an antibody that binds to at least one FucT or XylT, and reduces the activity of the FucT or XylT. In another embodiment, the binding of the antibody results in increased turnover of the antibody-FucT or XylT complex by cellular quality control  
30 mechanisms. The expression of antibodies in plant cells and the inhibition of molecular pathways by expression and binding of antibodies to proteins in plant cells are well known in the art. See, for example, Conrad and Sonnewald (2003) *Nature Biotech.* 21:35-36, incorporated herein by reference.

### Gene Disruption

In some embodiments of the present invention, the activity of FucT or XylT, or both, is reduced or eliminated by disrupting the gene encoding the FucT or XylT, or both. The gene encoding the FucT or XylT, or both, may be disrupted by any method known in the art. For example, in one embodiment, the gene is disrupted by transposon tagging. In another embodiment, the gene is disrupted by mutagenizing plants using random or targeted mutagenesis, and selecting for plants that have reduced FucT and/or XylT activity.

### Transposon Tagging

In one embodiment of the invention, transposon tagging is used to reduce or eliminate the activity of FucT or XylT, or both. Transposon tagging comprises inserting a transposon within an endogenous FucT or XylT gene to reduce or eliminate expression of the FucT or XylT. "FucT" or "XylT" gene is intended to mean the gene that encodes a FucT or XylT, respectively, according to the invention.

In this embodiment, the expression of FucT or XylT is reduced or eliminated by inserting a transposon within a regulatory region or coding region of the gene encoding the FucT or XylT. A transposon that is within an exon, intron, 5' or 3' untranslated sequence, a promoter, or any other regulatory sequence of a FucT or XylT, or both, gene may be used to reduce or eliminate the expression and/or activity of the encoded FucT or XylT.

Methods for the transposon tagging of specific genes in plants are well known in the art. See, for example, Maes *et al.* (1999) *Trends Plant Sci.* 4:90-96; Dharmapuri and Sonti (1999) *FEMS Microbiol. Lett.* 179:53-59; Meissner *et al.* (2000) *Plant J.* 22:265-274; Phogat *et al.* (2000) *J. Biosci.* 25:57-63; Walbot (2000) *Curr. Opin. Plant Biol.* 2:103-107; Gai *et al.* (2000) *Nucleic Acids Res.* 28:94-96; Fitzmaurice *et al.* (1999) *Genetics* 153:1919-1928). In addition, the TUSC process for selecting Mu insertions in selected genes has been described in Bensen *et al.* (1995) *Plant Cell* 7:75-84; Mena *et al.* (1996) *Science* 274:1537-1540; each of which is herein incorporated by reference.

The invention encompasses additional methods for reducing or eliminating the activity of FucT or XylT. Examples of other methods for altering or mutating a

genomic nucleotide sequence in a plant are known in the art and include, but are not limited to, the use of RNA:DNA vectors, RNA:DNA mutational vectors, RNA:DNA repair vectors, mixed-duplex oligonucleotides, self-complementary RNA:DNA oligonucleotides, and recombinogenic oligonucleobases. Such vectors and methods of use are known in the art. See, for example, U.S. Patent Nos. 5,565,350; 5,731,181; 5,756,325; 5,760,012; 5,795,972; and 5,871,984; each of which are herein incorporated by reference. See also, WO 98/49350, WO 99/07865, WO 99/25821, and Beetham *et al.* (1999) *Proc. Natl. Acad. Sci. USA* 96:8774-8778; each of which is herein incorporated by reference.

Thus inhibition of expression of FucT and/or XylT in a higher plant of interest can be accomplished by any of the foregoing methods in order to alter the *N*-glycosylation pattern of endogenous and heterologous glycoproteins produced within that plant such that these glycoproteins comprise complex *N*-glycans that have a reduction in the amount of  $\beta$ 1,2-linked xylose residues and/or  $\alpha$ 1,3-linked fucose residues. The extent to which attachment of the  $\beta$ 1,2-linked xylose residue and/or  $\alpha$ 1,3-linked fucose residue to glycoprotein *N*-glycans is reduced is governed by the degree of inhibition of expression of the respective XylT and FucT enzymes.

In some embodiments of the invention, recombinant glycoproteins produced in a plant host that is stably transformed using the methods described herein to target XylT expression have *N*-linked glycans comprising less than 50%, less than 40%, less than 30% of the  $\beta$ 1,2-linked xylose residues occurring in the respective *N*-linked glycans of glycoproteins produced in a plant host that has not been genetically modified to inhibit expression of the XylT enzyme and isoforms thereof. In other embodiments, these recombinant glycoproteins have *N*-linked glycans comprising less than 25%, less than 20%, less than 15%, less than 10%, less than 5%, or less than 1% of the  $\beta$ 1,2-linked xylose residues occurring in the respective *N*-linked glycans of glycoproteins produced in a plant host that has not been genetically modified to inhibit expression of the XylT enzyme and isoforms thereof. In yet other embodiments, the methods of the invention provide for complete silencing of the XylT gene and any isoforms thereof within the stably transformed plant, such that the recombinant glycoproteins produced within the plant have *N*-linked glycans that are devoid of  $\beta$ 1,2-linked xylose residues.



In like manner, where a plant host has been stably transformed using the methods described herein to target FucT expression, recombinant glycoproteins produced within the plant have *N*-linked glycans comprising less than 50%, less than 40%, less than 30% of the  $\alpha$ 1,3-linked fucose residues occurring in the respective *N*-linked glycans of glycoproteins produced in a plant host that has not been genetically modified to inhibit expression of the FucT enzyme and isoforms thereof. In other embodiments, these recombinant glycoproteins have *N*-linked glycans comprising less than 25%, less than 20%, less than 15%, less than 10%, less than 5%, or less than 1% of the  $\alpha$ 1,3-linked fucose residues occurring in the respective *N*-linked glycans of glycoproteins produced in a plant host that has not been genetically modified to inhibit expression of the FucT enzyme and isoforms thereof. In yet other embodiments, the methods of the invention provide for complete silencing of the FucT gene and any isoforms thereof within the stably transformed plant, such that the recombinant glycoproteins produced within the plant have *N*-linked glycans that are devoid of  $\alpha$ 1,3-linked fucose residues.

Where a plant host has been stably transformed using the methods described herein to target expression of both the XylT and FucT enzymes, and any isoforms thereof, recombinant glycoproteins produced within the plant have *N*-linked glycans comprising less than 50%, less than 40%, less than 30% of the  $\beta$ 1,2-linked xylose residues and less than 50%, less than 40%, less than 30% of the  $\alpha$ 1,3-linked fucose residues occurring in the respective *N*-linked glycans of glycoproteins produced in a plant host that has not been genetically modified to inhibit expression of the XylT and FucT enzymes and isoforms thereof. In other embodiments, these recombinant glycoproteins have *N*-linked glycans comprising less than 25%, less than 20%, less than 15%, less than 10%, less than 5%, or less than 1% of the  $\beta$ 1,2-linked xylose residues and 25%, less than 20%, less than 15%, less than 10%, less than 5%, or less than 1% of the  $\alpha$ 1,3-linked fucose residues occurring in the respective *N*-linked glycans of glycoproteins produced in a plant host that has not been genetically modified to inhibit expression of the XylT and FucT enzymes and isoforms thereof. In yet other embodiments, the methods of the invention provide for complete silencing of the XylT and FucT gene and any isoforms thereof within the stably transformed plant, such that the recombinant glycoproteins produced within the plant

have *N*-linked glycans that are devoid of  $\beta$ 1,2-linked xylose residues and  $\alpha$ 1,3-linked fucose residues.

5 In some embodiments of the present invention, a plant host that has been stably transformed using the methods described herein to target expression of both the XylT and FucT enzymes, and any isoforms thereof, is capable of producing recombinant glycoproteins wherein the *N*-linked glycans are substantially homogenous. By “substantially homogenous” is intended that the glycosylation profile reflects the presence of a single major peak corresponding to a desired *N*-glycan species, more particularly, the G0 glycan species, wherein at least 90% of the  
10 *N*-glycan structures present on said glycoproteins are of the G0 glycan species.

Methods for monitoring changes in the *N*-glycosylation pattern of glycoproteins, also referred to as glycosylation profiles, are well known in the art and include, but are not limited to, matrix-assisted laser desorption ionization time-of-flight (MALDI-TOF) mass spectrometry, for example, using the modified MALDI-  
15 TOF assay disclosed in Example 3 herein below, liquid chromatograph mass spectrometry (LC-MS), gas chromatography, anion-exchange chromatography, size-exclusion chromatography, high-concentration polyacrylamide gel electrophoresis, nuclear magnetic resonance spectroscopy, and capillary electrophoresis and capillary gel electrophoresis, fluorescence labeling and detection by high-performance liquid  
20 chromatography (HPLC) and QTOF, and the like. In this manner, changes in the *N*-glycosylation pattern due to inhibition of the expression or function of XylT and/or FucT in a stably transformed plant of the invention can be monitored by subjecting a sample (for example, a leaf tissue sample) obtained from the stably transformed plant to total *N*-glycan analysis by MALDI-TOF mass spectrometry, and comparing the  
25 results with those obtained for a comparable sample from a control plant, wherein the control plant has not been genetically modified to inhibit expression or function of XylT and/or FucT. A reduction in the amount of xylose and/or fucose residues in the *N*-glycans can be monitored by a reduction of the mass of the respective peaks. See, for example, Strasser *et al.* (2004) *FEBS Letters* 561:132-136.

30 Similarly, the glycosylation profile of any given recombinantly produced glycoprotein is readily determined using standard techniques well known to those in the art. See, for example, the review provided in Morelle and Michalski (2005) *Curr. Anal. Chem.* 1:29-57; herein incorporated by reference in its entirety. Thus,

glycoproteins that have been recombinantly produced in a host organism, including a plant host, can be analyzed for the ratio of the particular *N*-linked glycan structures attached thereto. In this manner, a sample comprising isolated recombinantly produced glycoprotein can be subjected to enzymatic or chemical reaction to release the individual glycan structures from the glycoprotein. Following this deglycosylation step, analysis of the glycosylation profile can be carried out using any of the analytical assays described herein above.

Recombinantly produced glycoprotein products typically exist as a diverse population of glycoforms carrying between one and several dozen different glycans in variable molar amounts at glycosylation sites with varying degrees of site occupancy. Depending upon the glycoprotein, different glycoforms can yield different functional profiles. Thus, in some embodiments of the invention, it is desirable to determine the glycosylation profile of the glycoprotein having the *N*-glycans intact. Any technique known in the art for determining the glycosylation profile of an intact glycoprotein can be used, including the mass spectrometry methods noted above and in the examples herein below.

By reducing or eliminating the expression or function of fucosyltransferase and/or xylosyltransferase in the manner set forth herein, either transiently or stably, it is possible to produce a transgenic higher plant having the ability to produce glycoproteins having an *N*-glycosylation profile with reduced heterogeneity relative to that normally observed for glycoproteins produced by this plant when expression or function of these enzymes has not been altered (i.e., the plant has the native or wild-type glycosylation machinery). Where expression or function of one or both of these enzymes is stably reduced or eliminated using one or more of the methods described herein above, the reduction in the heterogeneity of the *N*-glycosylation profile of glycoproteins produced by the transgenic plant can be maintained from plant generation to plant generation, including with asexual or sexual reproduction, and can be maintained across cultural conditions and with scale-up in production.

In this manner, the present invention provides a method for reducing heterogeneity of the *N*-glycosylation profile of a glycoprotein produced in a higher plant, for example, a dicotyledonous or monocotyledonous plant, for example, a duckweed plant. The method comprises introducing into the plant a nucleotide construct described herein such that the expression or function of fucosyltransferase

and/or xylosyltransferase is reduced or eliminated within the plant. In some embodiments of the invention, the method for reducing heterogeneity of the *N*-glycosylation profile of a glycoprotein produced in a higher plant comprises introducing into the higher plant of interest at least one nucleotide construct described  
5 herein above, where the nucleotide construct(s) provides for suppression of the expression of fucosyltransferase and/or xylosyltransferase in the plant, for example, using one or more of the methods described herein above.

By “reducing heterogeneity of the *N*-glycosylation profile” it is intended that the *N*-glycosylation profile is characterized by a reduction in the total number of  
10 distinct *N*-glycan species that appear in the profile. Thus, for example, where a glycoprotein produced by a higher plant having the native or wild-type glycosylation machinery (and thus which has not been genetically modified to reduce or eliminate expression of fucosyltransferase and xylosyltransferase) produces a glycoprotein with an *N*-glycosylation profile characterized by the presence of a mixture of 5 *N*-glycan  
15 species, the methods of the invention can be used to reduce the number of *N*-glycan species appearing in the *N*-glycosylation profile. In this manner, when that higher plant is genetically modified in the manner set forth herein to reduce or eliminate expression or function of fucosyltransferase and/or xylosyltransferase, the *N*-glycosylation profile of this glycoprotein would be characterized by a reduction in the  
20 number of *N*-glycan species appearing in the profile, for example, a mixture of fewer than 5 *N*-glycan species, for example, 4, 3, or 2 *N*-glycan species, or even a single *N*-glycan species. Where heterogeneity of the *N*-glycosylation profile is reduced such that the profile is characterized by the presence of a single predominant *N*-glycan species, the *N*-glycosylation profile would be substantially homogeneous for that *N*-  
25 glycan species.

In some embodiments, the methods for reducing the heterogeneity of the *N*-glycosylation profile of a glycoprotein produced in a higher plant result in the produced glycoprotein having an *N*-glycosylation profile that is substantially homogeneous for the G0 glycan species. In such embodiments, the methods for  
30 reducing the heterogeneity of the *N*-glycosylation profile of a glycoprotein produced in a higher plant result in the produced glycoprotein having a substantially homogeneous *N*-glycosylation profile, wherein at least 80%, at least 85%, at least 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, or at least 99% of the total

amount of *N*-glycan species appearing in the *N*-glycosylation profile for the glycoprotein is represented by the G0 glycan species. In these embodiments, a trace amount of precursor *N*-glycan species may appear in the *N*-glycosylation profile as noted elsewhere herein, where any given precursor *N*-glycan species that is present in the *N*-glycosylation profile is present at less than 5%, preferably less than 4%, less than 3%, less than 2%, less than 1%, and even less than 0.5% or even less than 0.1% of the total amount of *N*-glycan species appearing in the profile.

The glycoprotein can be an endogenous glycoprotein of interest, or can be a heterologous glycoprotein that is produced by the higher plant of interest, for example, a mammalian glycoprotein, including, for example, the glycoproteins described elsewhere herein. In some embodiments, the glycoprotein is a monoclonal antibody. In other embodiments, the glycoprotein is selected from the group consisting of an interferon, erythropoietin (EPO), tissue plasminogen activator (tPA), plasminogen, granulocyte-macrophage colony stimulating factor (GM-CSF), and therapeutic immunoglobulins.

Using the methods of the present invention, it is possible to maintain the reduced heterogeneity within the *N*-glycosylation profile for a glycoprotein produced in the transgenic plant with scale-up in production, and thus the plant continues to produce the glycoprotein such that its *N*-glycosylation profile is characterized by a reduction in the number of *N*-glycan species appearing in the profile. By “scale-up in production” or “increase in production scale” is intended an increase in the amount of plant biomass that is present within a culture system (i.e., a culture vessel or culture container within which the plant is cultured) that is being used to produce a protein of interest, in this case, a glycoprotein of interest. Thus, scale-up in production occurs, for example, when scaling up production from a scale that is suitable for research purposes to one that is suitable for pilot production, and further up to a scale that is suitable for commercial production of the glycoprotein of interest.

In some embodiments, the transgenic higher plant is a monocotyledonous plant, for example, a duckweed plant, that serves as a host for recombinant production of a glycoprotein, and the reduced heterogeneity of the *N*-glycosylation profile of the recombinantly produced glycoprotein is maintained with an increase in production scale, where the production scale is increased by at least 300-fold, at least 500-fold, at least 700-fold, at least 1,000-fold, at least 1,500-fold or greater over the initial starting



biomass. In some of these embodiments, the transgenic higher plant is a duckweed plant that recombinantly produces a glycoprotein of interest, and the reduced heterogeneity of the *N*-glycosylation profile is maintained with an increase in production scale, where the production scale is increased by at least 2,000-fold, at least 3,000-fold, at least 4,000-fold, at least 5,000-fold, at least 6,000-fold, at least 6,500-fold, or greater over the initial starting biomass. In one such embodiment, the higher plant is a duckweed plant that recombinantly produces a glycoprotein of interest, and the reduced heterogeneity of the *N*-glycosylation profile for that glycoprotein is maintained with an increase in production scale, where the production scale is increased by at least 7,000-fold, 8,000-fold, 9,000-fold, 10,000-fold, 12,500-fold, 15,000-fold, 17,500-fold, 20,000-fold, 23,000-fold, 26,000-fold, or greater over the initial starting biomass.

Furthermore, when the transgenic plant of interest is to be maintained by continuous clonal culture, the resulting transgenic line continues to produce glycoproteins that exhibit the reduced heterogeneity within their *N*-glycosylation profile. Continuous clonal culture can be achieved using any suitable method known in the art. In some embodiments, continuous clonal culture is achieved by periodically taking one or more subsamples of the plant culture and transferring the subsample(s) to fresh culture medium for further culture. Thus, for example, in some embodiments, the transgenic plant line that is maintained by continuous clonal culture is a duckweed transgenic plant line that has been genetically modified to reduce or eliminate expression or function of fucosyltransferase and/or xylosyltransferase. In this manner, the reduced heterogeneity of the *N*-glycosylation profile of glycoproteins produced in the transgenic plant line is maintained with continuous clonal culture of the transgenic plant line for at least 8 months, at least 10 months, at least 1 year, at least 1.5 years, at least 2 years, at least 2.5 years, at least 3 years, at least 3.5 years, at least 4 years, at least 4.5 years, at least 5 years, or longer, and can be maintained for as long as the transgenic plant line is maintained.

### Heterologous Polypeptides and Glycoproteins

Higher plants, particularly higher plants that serve as expression systems for recombinant proteins for pharmaceutical use, that have been stably transformed to produce glycoproteins with an altered *N*-glycosylation pattern using the methods

described herein may be genetically modified to produce any recombinant protein of interest. Where the recombinant protein is one in which post-translational glycosylation is applicable, the methods of the invention advantageously provide a means to produce these glycoproteins with an *N*-glycosylation pattern that more closely reflects that of mammalian hosts, particularly a glycosylation pattern that is “humanized.” Examples of recombinant proteins of interest include, but are not limited to, insulin, growth hormone, plasminogen,  $\alpha$ -interferon,  $\beta$ -interferon,  $\beta$ -glucocerebrosidase,  $\beta$ -glucuronidase, retinoblastoma protein, p53 protein, angiostatin, leptin, monoclonal antibodies and fragments thereof, cytokines, for example; erythropoietin (EPO), granulocyte macrophage colony stimulating factor, tissue plasminogen activator, blood coagulation factors, for example, Factor VII, Factor VIII, Factor IX, and activated protein C, alpha 1-antitrypsin, receptors, hormones, human vaccines, animal vaccines, peptides, and serum albumin. Furthermore, the transgenic higher plants of the invention are capable of producing a glycoprotein product that has a substantially homogenous glycosylation profile for the G0 glycan species, and which is characterized by its substantial homogeneity for the G0 glycoform. This advantageously results in plant host expression systems that have increased production consistency, as well as reduced chemical, manufacturing, and control (CMC) risk associated with the production of these glycoprotein compositions.

The glycoprotein compositions produced in accordance with the methods of the present invention may be contained in a composition comprising a pharmaceutically acceptable carrier. Such compositions are useful in a method of treating a subject for a disease or disorder for which treatment with the glycoprotein will provide a therapeutic benefit. In this manner, glycoproteins that have been produced in a plant, for example, a duckweed plant, stably transformed in accordance with the methods of the present invention can be administered to a subject in need thereof.

The glycoprotein compositions of the invention comprise *N*-linked glycans that are predominately of the G0 glycan structure. In this manner, the present invention provides glycoprotein compositions that have glycosylation profiles that are “substantially homogeneous” or “substantially uniform” or have “substantial homogeneity” as defined herein above. Thus, in some embodiments, the glycoprotein compositions are substantially homogeneous for the G0 glycan species, and thus have

a substantially homogeneous glycosylation profile, wherein at least 80%, at least 85%, at least 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, or at least 99% of the total amount of *N*-glycan species appearing in the glycosylation profile for the composition is represented by the G0 glycan species, with a trace amount of precursor *N*-glycan species appearing in the glycosylation profile, i.e., any given precursor *N*-glycan species that is present in the glycosylation profile is present at less than 5%, preferably less than 4%, less than 3%, less than 2%, less than 1%, and even less than 0.5% or even less than 0.1% of the total amount of *N*-glycan species appearing in the profile. For such a composition, a representative precursor *N*-glycan species appearing in its glycosylation profile would be the Man3GlcNAc2, MGn (GlcNAc1Man3GlcNAc2 wherein GlcNAc1 is attached to the 1,3 mannose arm), and GnM (GlcNAc1Man3GlcNAc2 wherein GlcNAc1 is attached to the 1,6 mannose arm) precursor *N*-glycan species described above, where any single one or any combination of these precursor *N*-glycan species can be present.

In this manner, the invention provides “substantially homogeneous” or “substantially uniform” glycoprotein compositions or glycoprotein compositions having “substantial homogeneity” as defined herein above. In some embodiments, the invention provides substantially homogeneous glycoprotein compositions, wherein at least 80%, at least 85%, at least 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, or at least 99% of the glycoprotein present in the composition is represented by the G0 glycoform, wherein all anticipated glycosylation sites are occupied by the G0 glycan species, with a trace amount of precursor or undesired glycoforms being present in the composition, i.e., the precursor glycoforms represent less than 5%, less than 4%, less than 3%, less than 2%, less than 1%, or even less than 0.5%, or less than 0.1% of the total glycoforms present within the composition. In such a composition, a representative precursor glycoform would be one in which glycosylation sites are unoccupied, and an exemplary undesired glycoform would be a glycoform having a mixture of G0 glycan and G0X or G0XF3 glycan species attached to its glycosylation sites.

In some embodiments of the invention, the plant host comprises one or more polynucleotides that provide for expression of an antibody that specifically binds to a mammalian protein of interest, particularly a human protein of interest. Thus, in one aspect, the invention provides methods for producing monoclonal antibodies in higher

plants, wherein the monoclonal antibodies have an *N*-glycosylation pattern that reflects a reduction in the amount of  $\beta$ 1,2-linked xylose residues and  $\alpha$ 1,3-linked fucose residues within the *N*-linked glycans, and compositions comprising recombinant monoclonal antibodies produced using plant hosts genetically modified in the manner set forth herein. In some embodiments, the plant host of interest is a member of the duckweed family.

Monoclonal antibodies are increasingly being used as therapeutic agents to treat human disease, including, but not limited to, cancer and diseases having an autoimmune or inflammatory component. See, for example, King (1999) *Curr. Opin. Drug Discovery Dev.* 2:110-17; Vaswani and Hamilton (1998) *Ann. Allergy Asthma Immunol.* 81:105-19; and Holliger and Hoogenboom, *Nat. Biotechnology* 16:1015-16; each of which is herein incorporated by reference. Although some of these antibodies have therapeutic effects that result solely from antigen binding, for example antibodies that bind to a receptor or ligand to prevent ligand-receptor interactions, other antibodies need effector functions such as the recruitment of the immune system to kill target cells in order to be therapeutically active. See, for example, Clynes *et al.* (2000) *Nat. Med.* 6:443-46; Clynes *et al.* (1998) *Proc. Natl. Acad. Sci. U.S.A.* 95:652-56, and Anderson *et al.* (1997) *Biochem. Soc. Trans.* 25:705-8; each of which is herein incorporated by reference.

The antigen-recognition activities and effector functions of antibodies reside in different portions of the antibody molecule. The Fab' portion of the antibody provides antigen recognition activity, while the Fc portion provides effector functions such as the activation of accessory effector cells including phagocytic cells (macrophages and neutrophils), natural killer cells, and mast cells. Antibodies bind to cells via the Fc region, with an Fc receptor site on the antibody Fc region binding to an Fc receptor (FcR) on a cell. There are a number of Fc receptors that are specific for different classes of antibodies, including IgG (gamma receptors), IgE (eta receptors), IgA (alpha receptors) and IgM (mu receptors). Binding of antibody to Fc receptors on cell surfaces triggers a number of important and diverse biological responses including engulfment and destruction of antibody-coated particles, clearance of immune complexes, lysis of antibody-coated target cells by killer cells (called antibody-dependent cell-mediated cytotoxicity, or ADCC), initiation of complement-dependent cytotoxicity (CDC), release of inflammatory mediators, and

control of immunoglobulin production. Methods for assaying effector function of antibodies are well known in the art and include those assaying for CDC, ADCC, and apoptosis. See, for example, Subbramanian *et al.* (2002) *J. Clin. Microbiol.* 40:2141-2146; Ahman *et al.* (1994) *J. Immunol. Methods* 36:243-254; Brezicka *et al.* (2000) *Cancer Immunol. Immunother.* 49:235-242; Gazzano-Santoro *et al.* (1997) *J. Immunol. Methods* 202:163-171; Prang *et al.* (2005) *British J. Cancer* 92:342-349; Shan *et al.* (1998) *Blood* 92:3756-3771; Ghetie *et al.* (2001) *Blood* 97:1392-1398; and, Mathas *et al.* (2000) *Cancer Research* 60:7170-7176; all of which are herein incorporated by reference.

It is known in the art that the glycosylation status of the Fc portion of an antibody molecule plays a key role in determining whether an antibody will have effector function. See, for example, Tao and Morrison (1987) *J. Immunol.* 143:2595-601; Wright and Morrison (1997) *Trends in Biotech.* 15:26-32; Wright and Morrison (1998) *J. Immunol.* 160:3393-402; Mimura *et al.* (2000) *Mol. Immunol.* 37:697-706; Jefferis and Lund (2002) *Immunol. Lett.* 82:57-65; Krapp *et al.* (2003) *J. Mol. Biol.* 325:979-89; and Jefferis (2005) *Biotechnol. Prog.* 21:11-16; each of which is herein incorporated by reference. Glycosylation of recombinantly produced antibodies varies depending on the expression system used. See, for example, Raju *et al.* (2000) *Glycobiology* 10:477-86; Wright and Morrison (1997) *Trends in Biotech.* 15:26-32. Further, where the *N*-glycosylation pattern of a mammalian-produced monoclonal antibody is altered to reduce or deplete the  $\alpha(1,6)$ -linked core fucose residue, the monoclonal antibody exhibits increased effector function, particularly increased ADCC activity. See, for example, U.S. Patent No. 6,946,292. For some mammalian-produced monoclonal antibodies, where the *N*-glycosylation pattern is altered to reduce or deplete the  $\beta(1,4)$ -galactose residues attached to the 1,3 and/or 1,6 mannose arms, activation of complement-dependent cytotoxicity (CDC) against antigen-bearing target cells may be reduced without altering other functional activities of the antibody, including ADCC activity. See, for example, Boyd *et al.* (1995) *Mol. Immunol.* 32:1311-1318.

Antibodies having antigen recognition activity, and in some embodiments improved effector function, may be produced by a higher plant host, such as duckweed, that has been stably transformed in the manner set forth herein to alter its glycosylation machinery. Accordingly, the present invention provides methods for



producing a recombinant monoclonal antibody, including a monoclonal antibody having improved effector function, wherein the antibody is recombinantly produced within a plant having an altered *N*-glycosylation pattern of endogenous and heterologous glycoproteins produced therein such that these glycoproteins exhibit a reduction in the amount of the plant-specific  $\beta$ 1,2-linked xylose residues and/or  $\alpha$ 1,3-linked fucose residues attached to the *N*-glycans thereof. Where the antibodies have reduced amounts  $\alpha$ 1,3-linked fucose residues attached to the *N*-glycans thereof, the antibodies may have increased ADCC activity relative to antibodies produced in a control plant that has not been genetically modified to inhibit expression or function of FucT.

Also encompassed are recombinant monoclonal antibodies having effector function, and in some embodiments, improved effector function, where the antibodies are produced in a duckweed expression system that has been genetically modified to inhibit expression or function of the FucT of SEQ ID NO:3 and/or the XylT of SEQ ID NO:6, and any isoforms thereof, for example, the XylT isoform #2 comprising the sequence set forth in SEQ ID NO:21 (encoded by SEQ ID NO:20). Thus, in some embodiments, the plant serving as the host for recombinant production of the monoclonal antibody is a member of the *Lemnaceae* as noted elsewhere herein, for example, a *Lemna* plant, comprising, for example, a XylT RNAi expression cassette and/or a FucT RNAi expression cassette described above stably integrated within its genome. In this manner, the present invention provides a method for producing a recombinant monoclonal antibody having an *N*-glycosylation pattern that more closely resembles that found in a mammalian host expression system, and with improved effector function, where the method comprises expressing one or more chains of the antibody in a duckweed plant, or duckweed cell or duckweed nodule, that has been genetically modified to alter the glycosylation machinery such that the recombinantly produced monoclonal antibody exhibits a reduction in the attachment of the plant  $\beta$ 1,2-linked xylose residue and/or  $\alpha$ 1,3-linked fucose residue to the *N*-glycans thereof, and culturing the duckweed plant, or duckweed cell or duckweed nodule, under conditions suitable for expression of the monoclonal antibody.

Thus the present invention provides novel antibody compositions wherein the antibody comprises *N*-linked glycans that are predominately of the G0 glycan structure. In this manner, the present invention provides antibody compositions, for

example, monoclonal antibody compositions, that have glycosylation profiles that are “substantially homogeneous” or “substantially uniform” or have “substantial homogeneity” as defined herein above. Thus, in some embodiments, the antibody compositions are substantially homogeneous for the G0 glycan species, and thus have a substantially homogeneous glycosylation profile, wherein at least 80%, at least 85%, at least 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, or at least 99% of the total amount of *N*-glycan species appearing in the glycosylation profile for the composition is represented by the G0 glycan species, with a trace amount of precursor *N*-glycan species appearing in the glycosylation profile, i.e., any given precursor *N*-glycan species that is present in the glycosylation profile is present at less than 5%, preferably less than 4%, less than 3%, less than 2%, less than 1%, and even less than 0.5% or even less than 0.1% of the total amount of *N*-glycan species appearing in the profile. For such a composition, a representative precursor *N*-glycan species appearing in its glycosylation profile would be, for example, the Man<sub>3</sub>GlcNAc<sub>2</sub>, MGn (GlcNAc1Man<sub>3</sub>GlcNAc<sub>2</sub> wherein GlcNAc1 is attached to the 1,3 mannose arm), and GnM (GlcNAc1Man<sub>3</sub>GlcNAc<sub>2</sub> wherein GlcNAc1 is attached to the 1,6 mannose arm) precursor *N*-glycan species described above, where any single one or any combination of these precursor *N*-glycan species can be present.

In one such embodiment, the monoclonal antibody composition has a substantially homogeneous glycosylation profile, wherein 95.8% of the total amount of *N*-glycan species appearing in the glycosylation profile for the composition is represented by the G0 glycan species (GlcNAc<sub>2</sub>Man<sub>3</sub>GlcNAc<sub>2</sub>), with the following precursor *N*-glycan species appearing in the glycosylation profile: Man<sub>3</sub>GlcNAc<sub>2</sub> (0.67%), GlcNAcMan<sub>3</sub>GlcNAc<sub>2</sub> (1.6%), GalGlcNAc<sub>2</sub>Man<sub>3</sub>GlcNAc<sub>2</sub> (1.2%), Man<sub>6</sub>GlcNAc<sub>2</sub> (0.21%), Man<sub>7</sub>GlcNAc<sub>2</sub> (0.30%), and Man<sub>8</sub>GlcNAc<sub>2</sub> (0.28%). This can be compared with the monoclonal antibody composition obtained from the “wild-type” duckweed plant expression system wherein the same monoclonal antibody is expressed but where the glycosylation machinery of the duckweed plant has not been genetically modified to inhibit expression of XylT and FucT. Such a “wild-type”-derived monoclonal antibody composition has a more heterogeneous glycosylation profile that is characterized by two predominant *N*-glycan species, i.e., G0XF<sup>3</sup> and G0X, with several precursor *N*-glycan species represented in trace amounts. In one such embodiment, the “wild-type”-derived monoclonal antibody composition has a

glycosylation profile with the following *N*-glycan species represented therein: G0 (GlcNAc<sub>2</sub>Man<sub>3</sub>GlcNAc<sub>2</sub>) (8.4%); G0X (GlcNAc<sub>2</sub>[Xyl]Man<sub>3</sub>GlcNAc<sub>2</sub>) (17.2%); G0XF<sup>3</sup> (GlcNAc<sub>2</sub>[Xyl]Man<sub>3</sub>[Fuc]GlcNAc<sub>2</sub>) (67.4%); Man<sub>3</sub>GlcNAc<sub>2</sub> (0.26%); GlcNAcMan<sub>3</sub>GlcNAc<sub>2</sub> (0.40%); (Xyl)Man<sub>3</sub>(Fuc)GlcNAc<sub>2</sub> (0.76%);  
 5 GlcNAc<sub>2</sub>Man<sub>3</sub>(Fuc)GlcNAc<sub>2</sub> (2.1%); GlcNAc(Xyl)Man<sub>3</sub>(Fuc)GlcNAc<sub>2</sub> (1.4%); Man<sub>6</sub>GlcNAc<sub>2</sub> (0.21%); Man<sub>7</sub>GlcNAc<sub>2</sub> (0.63%); Gal(Fuc)GlcNAc<sub>2</sub>(Xyl)Man<sub>3</sub>(Fuc)GlcNAc<sub>2</sub> (0.26%); Man<sub>8</sub>GlcNAc<sub>2</sub> (0.61%); and Man<sub>9</sub>GlcNAc<sub>2</sub> (0.40%).

In this manner, the invention provides “substantially homogeneous” or  
 10 “substantially uniform” antibody compositions or antibody compositions having “substantial homogeneity” as defined herein above. In some embodiments, the invention provides substantially homogeneous antibody compositions, for example, monoclonal antibody compositions, wherein at least 80%, at least 85%, at least 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, or at least 99% of the antibody present  
 15 in the composition is represented by the G0 glycoform, wherein all anticipated glycosylation sites (for example, each of the Asn-297 residues of the C<sub>H</sub>2 domains of the heavy chains of an IgG-type antibody) are occupied by the G0 glycan species, with a trace amount of precursor glycoforms being present in the composition. In one such composition, the precursor glycoforms are selected from the group consisting of  
 20 an antibody having an Fc region wherein the C<sub>H</sub>2 domain of one heavy chain has a G0 glycan species attached to Asn 297, and the C<sub>H</sub>2 domain of the other heavy chain is unglycosylated; an antibody having an Fc region wherein the C<sub>H</sub>2 domain of one heavy chain has a G0 glycan species attached to Asn 297, and the C<sub>H</sub>2 domain of the other heavy chain has the GnM or MGn precursor glycan attached to Asn 297; and an  
 25 antibody having an Fc region wherein the Asn 297 glycosylation site on each of the C<sub>H</sub>2 domains has a G0 glycan species attached, with a third G0 glycan species attached to an additional glycosylation site within the mAb structure; wherein a trace amount of these precursor glycoforms is present, i.e., the precursor glycoforms represent less than 5%, less than 4%, less than 3%, less than 2%, less than 1%, or  
 30 even less than 0.5%, or less than 0.1% of the total glycoforms present within the antibody composition.

The substantially homogeneous antibody compositions of the invention wherein at least 80%, at least 85%, at least 90%, 91%, 92%, 93%, 94%, 95%, 96%,

97%, 98%, or at least 99% of the antibody present in the composition is represented by the G0 glycoform represent “glycan-optimized antibodies.” By “glycan-optimized antibodies” is intended the antibodies of the invention have been genetically engineered in their glycosylation pattern such that they are substantially homogeneous for the G0 glycoform, which yields an antibody having improved Fc effector function. By “improved Fc effector function” is intended these antibodies have increased ADCC activity relative to same-sequence antibodies (i.e., antibodies that have the same amino acid sequence) that, as a result of the production process, have a more heterogeneous glycosylation profile. Thus, for example, antibodies produced in mammalian host expression systems, for example CHO cells, in insect host cells, in yeast cells, or in other plant host expression systems that have not been genetically altered to inhibit XylT and FucT expression tend to have more heterogeneous glycosylation profiles, and thus a mixture of glycoforms, that can effect overall effector function of the antibody product. The G0 glycoform of the antibody compositions of the present invention advantageously provides an antibody composition that has increased ADCC activity in association with the absence of fucose residues. In some embodiments, ADCC activity is increased by 25-fold, 50-fold, 75-fold, 100-fold, 150-fold, 200-fold, 250-fold, 300-fold, 400-fold, 500-fold, or even 1000-fold relative to same-sequence antibodies having a heterogeneous glycosylation profile (i.e., with multiple glycoforms present as major glycoforms in the antibody composition). Furthermore, the G0 glycoform lacks the terminal Gal residues present in antibodies having the G2 glycoform. As such, these substantially homogeneous antibody compositions of the invention having predominately the G0 glycoform have increased ADCC/CDC ratios. In addition, the substantially homogeneous antibody compositions of the invention having predominately the G0 glycoform have increased binding to the FcγRIII, for example, FcγRIIIa, wherein binding affinity is increased about 20-fold, 30-fold, 40-fold, 50-fold, 75-fold, up to 100-fold over that observed for same-sequence antibody compositions having a heterogeneous glycosylation profile, and thus a mixture of glycoforms. For oncology and autoimmune diseases, therapeutic antibodies having increased binding affinity for Fc receptors, for example, FcγRIII, has been strongly correlated with increased efficacy and improved response to treatment.

In some embodiments of the invention, the substantially homogeneous antibody compositions of the invention having predominately the G0 glycoform have altered CDC activity when compared to that observed for same-sequence antibody compositions having a heterogeneous glycosylation profile, and thus a mixture of glycoforms. For example, in one such embodiment, the substantially homogeneous antibody compositions of the invention have predominately the G0 glycoform and have decreased CDC activity when compared to that observed for same-sequence antibody compositions having a heterogeneous glycosylation profile, and thus a mixture of glycoforms. Thus, in some embodiments, the present invention provides substantially homogeneous antibody compositions having predominately the G0 glycoform and CDC activity that is decreased by as much as 5%, 10%, 15%, 20%, 25%, 30%, 35%, 40%, 45%, 50%, 55%, 60%, 65%, 70%, 75%, 80%, 85%, 90%, 95%, or even 100% when compared to same-sequence antibody compositions having a heterogeneous glycosylation profile. In some of these embodiments, the substantially homogeneous antibody compositions having predominately the G0 glycoform and decreased CDC activity are further characterized by having increased ADCC activity when compared to same-sequence antibody compositions having a heterogeneous glycosylation profile.

Without being bound by any theory or mechanism of action, a substantially homogeneous G0 glycoform antibody composition of the present invention having decreased CDC activity, and similar or increased ADCC activity in the manner described above, may advantageously provide for increased cytotoxicity against target cells while reducing potential adverse side effects that may be associated with complement activation following its administration. By reducing the potential for these adverse side effects, the substantially homogeneous antibody compositions having predominately the G0 glycoform can advantageously be administered at faster infusion rates, thereby reducing dosing time at any given administration, and/or can be dosed at higher initial concentrations if warranted, with reduced concern for triggering adverse side effects associated with complement activation.

For example, complement activation plays a pivotal role in the pathogenesis of moderate to severe first-dose side effects of treatment with the chimeric anti-CD20 monoclonal antibody IDEC-C2B8 (IDEC Pharmaceuticals Corp., San Diego, California; commercially available under the tradename Rituxan®, also referred to as



rituximab). See, for example, van der Kolk *et al.* (2001) *British J. Haematol.* 115:807-811. The rituximab antibody within the Rituxan® product is expressed within Chinese hamster ovary (CHO) cells, and thus the antibody composition comprises a heterogeneous glycosylation profile (i.e., a mixture of glycoforms). CDC activity of rituximab has been shown to be correlated with galactose content. In this manner, as the number of galactose residues increases from 0-2 moles/mole of heavy chain, the level of CDC activity increases from 80% ( $\beta$ -galactosidase treated to remove all  $\beta$ (1,4)-galactose residues from the 1,3 and 1,6 mannose arms of the *N*-glycans attached to Asn 297 of the C<sub>H</sub>2 domains of the heavy chains) to 150% (UDP galactosyl transferase treated to ensure  $\beta$ (1,4)-galactose residues are attached to both the 1,3 and 1,6 mannose arms of the *N*-glycans attached to Asn 297 sites) of the maximum observed for the antibody having 1 mole galactose/mole of heavy chain (see, IDEC BLA 97-0260 at the website [fda.gov/Cder/biologics/review/ritugen112697](http://fda.gov/Cder/biologics/review/ritugen112697), available on the worldwide web). A substantially homogeneous anti-CD20 antibody composition comprising anti-CD20 antibody having the same sequence as rituximab and having predominately the G0 glycoform would advantageously have decreased CDC activity, thereby reducing the potential for adverse side effects normally associated with complement activation upon antibody administration when the antibody composition comprises a heterogeneous glycosylation profile (i.e., a mixture of glycoforms).

Thus, a substantially homogeneous G0 glycoform antibody composition of the present invention having decreased CDC activity, and the same or increased ADCC activity, can advantageously be used in therapeutic applications that have heretofore been unsuitable, inadvisable, or inefficacious for one or more patient populations as a result of complications due to adverse side effects normally associated with complement activation upon administration of the same-sequence antibody composition that comprises a heterogeneous glycosylation profile (i.e., a mixture of glycoforms). Such side effects include, but are not limited to, moderate to severe side effects that can be associated with first-time and/or rapid administration of an antibody, including, for example, fever and/or chills, nausea, dyspnea, flushes, and the like. See, for example, van der Kolk *et al.* (2001) *British J. Haematol.* 115:807-811 and the references cited therein; Winkler *et al.* (1999) *Blood* 94:2217-2224). In this manner, the present invention provides a method for reducing one or more adverse

side effects related to complement activation upon administration of an antibody, for example, a monoclonal antibody, the method comprising administering the antibody as a substantially homogeneous antibody composition as defined herein above, and thus at least 80%, at least 85%, at least 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 5 98%, or at least 99% of the antibody present in the composition is represented by the G0 glycoform, with a trace amount of precursor glycoforms being present in the composition. In some embodiments, at least 90%, at least 95%, or at least 99% of the antibody present in the composition is represented by the G0 glycoform, with a trace amount of precursor glycoforms being present in the composition.

10 As a result of their increased Fc effector function, the substantially homogenous antibody compositions of the invention having the predominately G0 glycoform provide the opportunities for new and improved routes of administration, for example, extending the possible routes of administration for known therapeutic antibodies to other routes of administration beyond infusion and intravenous 15 administration, for example, to subcutaneous administration. Furthermore, as a result of their increased potency, the antibody compositions of the invention can be dosed at lower concentrations, or dosed at lower volumes, and dosed with less frequency. A reduction in the volume of the administered antibody composition is particularly advantageous in those instances where adverse events resulting from infusion 20 reactions with a monoclonal antibody are volume-related. The increased potency of the antibody compositions of the invention also opens up new clinical indications for existing mAb targets that may not have been responsive to antibody therapy with more heterogenous glycoform antibody compositions, and provides the ability to revisit previously undeveloped mAb targets.

25 The monoclonal antibodies produced in accordance with the methods of the present invention may be contained in a composition comprising a pharmaceutically acceptable carrier. Such compositions are useful in a method of treating a subject in need of an antibody having effector function, and in some embodiments, improved effector function where FucT expression has been targeted for inhibition. In this 30 manner, monoclonal antibodies produced in a plant, for example, a duckweed plant, stably transformed in accordance with the methods of the present invention can be administered to a subject in need thereof.

Any antibody may be produced within a plant host genetically modified according to the methods of the invention. In one embodiment, the antibody is a therapeutic antibody. Antibodies have been approved for the treatment of a number of disorders, and many additional therapeutic antibodies are in development. See,

5 Brekke and Sandlie (2003) *Nat. Rev. Drug. Discov.* 2(3):240, herein incorporated by reference. Examples of therapeutic antibodies that may be expressed by the methods of the invention include, but are not limited to, anti-CD3 antibodies targeting the CD3 antigen (*e.g.*, OKT®3); anti-platelet gpIIb/IIIa antibodies (*e.g.* abciximab (previously known as c7E3 Fab), distributed as ReoPro®); anti-EpCAM antibodies (*e.g.*,

10 Panorex®); anti-CD20 antibodies targeting the CD20 antigen (*e.g.* rituximab, distributed as Rituxan®; see U.S. Patent No. 5,736,137, herein incorporated by reference); anti IL-2 receptor (*e.g.* Zenapax®, Simulect®); anti-ERBB2 (*e.g.* trastuzumab, distributed as Herceptin®; see U.S. Patent No. 6,165,464, herein incorporated by reference); anti-TNF- $\alpha$  (*e.g.*, infliximab, distributed as Remicade®),

15 anti-F-protein (*e.g.* Synagis®); anti-CD30 antibodies targeting the CD30 antigen (see, for example, the 5F11 antibody described in Borchmann *et al.* (2003) *Blood* 102:3737-3742 and in Example 6 herein below, see also WO 03/059282 and U.S. Patent Application Publication No. 2004/0006215, herein incorporated by reference; the SGN-30 antibody described in Wahl *et al.* (2002) *Cancer Res.* 62:3736-3742, a

20 chimeric version of the AC10 anti-CD30 antibody, see also U.S. Patent Application Publication No. 20040018194 and U.S. Patent No. 7,090,843, herein incorporated by reference); anti-CD33 antibodies targeting the CD33 antigen (*e.g.* Mylotarg®; see U.S. Patent No. 5,733,001, herein incorporated by reference); anti-CD52 antibodies targeting the CD52 antigen (*e.g.*, alemtuzumab, distributed as Campath®; see U.S.

25 Patent No. 5,846,534, herein incorporated by reference); anti-CD20 (*e.g.*, Zevalin®); anti-IgE Fc (*e.g.*, omalizumab, marketed under Xolair®); anti-VEGF (*e.g.*, bevacizumab, distributed as Avastin®; see, Presta *et al.* (1997) *Cancer Res.* 57:4593-9; herein incorporated by reference); anti-EGF-R (*e.g.*, cetuximab, distributed as Erbitux®); anti-CD11a (*e.g.*, efalizumab, distributed under Raptiva®), anti-IL-8; anti-C5; anti-TNF- $\alpha$  (*e.g.*, adalimumab, distributed as Humira®); anti-TGF- $\beta$ 2; anti-

30 IL2 receptor (*e.g.*, daclizumab (*Zenapax*) and basiliximab (*Simulect*)); anti- $\alpha$ -4-integrin; anti-CD4; anti-CD2; anti-CD19 (*e.g.*, MT103, a bispecific antibody); anti-CD22 antibody targeting the CD22 antigen (*e.g.*, the monoclonal antibody BL-22);

anti-CD23 antibody targeting the CD23 antigen on malignant B cells (*e.g.*, IDEC-152); anti-CD80 antibody targeting the CD80 antigen (*e.g.*, IDEC-114); anti-CD38 antibody targeting the CD38 antigen on malignant B cells;  $\alpha$ -M-CSF antibody targeting macrophage colony stimulating factor; antibodies targeting the receptor activator of nuclear factor-kappaB (RANK) and its ligand (RANKL), which are overexpressed in multiple myeloma; anti-CD40 antibodies (*e.g.*, SGN-40) targeting the CD40 antigen on malignant B cells; antibodies targeting tumor necrosis factor-related apoptosis-inducing ligand receptor 1 (TRAIL-R1) (*e.g.*, the agonistic human monoclonal antibody HGS-ETR1) and TRAIL-R2 expressed on a number of solid tumors and tumors of hematopoietic origin) anti-HBV antibodies; or anti-MHC class II antibodies. Many of these antibodies are known to require effector activity for function, either CDC and/or ADCC, for example, Herceptin®, Rituxan®, Mylotarg®, and Campath®.

In some embodiments, the plant host that has been genetically modified to alter its glycosylation machinery in the manner set forth herein, for example, duckweed, may express biologically active  $\alpha$ -2b-interferon, for example human  $\alpha$ -2b-interferon precursor (NCBI Protein Accession No. AAB59402) or mature human  $\alpha$ -2b-interferon (amino acids 24-188 of NCBI Protein Accession No. AAB9402) or biologically active variants thereof. Examples of biologically active variants of human  $\alpha$ -2b-interferon are known in the art. See, for example, WO 2005/035767, disclosing truncated versions of  $\alpha$ -2b-interferon; European patent EP211148B1; and U.S. Patent Nos. 4,748,233, 4,801,685, 4,816,566, 4,973,479, 4,975,276, 5,089,400, 5,098,703, 5,231,176, and 5,869,293; herein incorporated by reference. In other embodiments, the plant host, for example, duckweed, may express biologically active mature human growth hormone with or without its accompanying signal peptide. The plant host may also express biologically active variants of human growth hormone. See, for example, WO 02/10414, directed to expression of biologically active polypeptides using a duckweed expression system.

In yet other embodiments, the plant host, for example, duckweed, may express plasminogen, microplasminogen, or biologically active variants thereof. The plasminogen or microplasminogen to be expressed in the plant host may be from any mammalian source. In some embodiments, the plasminogen or microplasminogen is

human or porcine. See, for example, WO 2005/078109, directed to expression of these proteins using a duckweed expression system.

By “biologically active variant” of these heterologous polypeptides is intended a polypeptide derived from the native polypeptide by deletion (so-called truncation) or addition of one or more amino acids to the N-terminal and/or C-terminal end of the native protein; deletion or addition of one or more amino acids at one or more sites in the native protein; or substitution of one or more amino acids at one or more sites in the native protein. Variants of a heterologous polypeptide of interest encompassed by the present invention retain the biological activity of the native polypeptides.

For example, biologically active variants of  $\alpha$ -2b-interferon continue to possess the desired biological activity of the native  $\alpha$ -2b-interferon including the ability to increase resistance to viral infection, or the ability to modulate the transcription of  $\alpha$ -2b-interferon-regulated gene targets. Such biologically active variants may result from, for example, genetic polymorphism or from human manipulation. Biologically active variants of a heterologous protein of interest will have at least about 50%, 60%, 65%, 70%, generally at least about 75%, 80%, 85%, preferably at least about 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, and more preferably at least about 98%, 99% or more sequence identity to the amino acid sequence for the heterologous protein of interest. Thus, for example, a biologically active variant of native  $\alpha$ -2b-interferon will have at least about 50%, 60%, 65%, 70%, generally at least about 75%, 80%, 85%, preferably at least about 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, and more preferably at least about 98%, 99% or more sequence identity to the amino acid sequence for human  $\alpha$ -2b-interferon precursor (NCBI Protein Accession No. AAB59402) or mature human  $\alpha$ -2b-interferon (amino acids 24-188 of NCBI Protein Accession No. AAB9402), where sequence identity is determined using the parameters identified herein above. Thus, a biologically active variant of a native heterologous protein of interest may differ from that protein by as few as 1-15 amino acid residues, as few as 1-10, such as 6-10, as few as 5, as few as 4, 3, 2, or even 1 amino acid residue.

In one embodiment, the plant host that has been genetically modified to alter its glycosylation machinery in the manner set forth herein is a stably transformed duckweed plant or duckweed cell or duckweed nodule that expresses biologically active polypeptides that cannot effectively be commercially produced by existing



gene expression systems, because of cost or logistical constraints, or both. For example, some proteins cannot be expressed in mammalian systems because the protein interferes with cell viability, cell proliferation, cellular differentiation, or protein assembly in mammalian cells. Such proteins include, but are not limited to, retinoblastoma protein, p53, angiostatin, and leptin. The present invention can be advantageously employed to produce mammalian regulatory proteins; it is unlikely given the large evolutionary distance between higher plants and mammals that these proteins will interfere with regulatory processes in duckweed. Transgenic duckweed can also be used to produce large quantities of proteins such as serum albumin (in particular, human serum albumin), hemoglobin, and collagen, which challenge the production capabilities of existing expression systems.

Finally, higher plant systems can be engineered to produce biologically active multimeric proteins (e.g., monoclonal antibodies, hemoglobin, P450 oxidase, and collagen, and the like) far more easily than can mammalian systems. One exemplary approach for producing biologically active multimeric proteins in duckweed uses an expression vector containing the genes encoding all of the polypeptide subunits. See, e.g., During *et al.* (1990) *Plant Mol. Biol.* 15:281 and van Engelen *et al.* (1994) *Plant Mol. Biol.* 26:1701. The expression cassette comprising the XylT and/or FucT inhibitory polynucleotide can be introduced into such a vector. This vector is then introduced into duckweed cells using any known transformation method, such as a ballistic bombardment or *Agrobacterium*-mediated transformation. This method results in clonal cell lines that express all of the polypeptides necessary to assemble the multimeric protein, as well as the XylT and/or FucT inhibitory sequences that alter the glycosylation pattern of the *N*-glycans of glycoproteins. Accordingly, in some embodiments, the transformed duckweed contains one or more expression vectors encoding a heavy and light chain of a monoclonal antibody or Fab' antibody fragment, and an expression cassette comprising the XylT and/or FucT inhibitory polynucleotide, and the monoclonal antibody or antibody fragment is assembled in the duckweed plant from the expressed heavy and light chain.

A variation on this approach is to make single gene constructs, mix DNA from these constructs together, then deliver this mixture of DNAs into plant cells using ballistic bombardment or *Agrobacterium*-mediated transformation. As a further variation, some or all of the vectors may encode more than one subunit of the

multimeric protein (i.e., so that there are fewer duckweed clones to be crossed than the number of subunits in the multimeric protein). In an alternative embodiment, each duckweed clone has been genetically modified to alter its glycosylation machinery and expresses at least one of the subunits of the multimeric protein, and duckweed clones secreting each subunit are cultured together and the multimeric protein is assembled in the media from the various secreted subunits. In some instances, it may be desirable to produce less than all of the subunits of a multimeric protein, or even a single protein subunit, in a transformed duckweed plant or duckweed nodule culture, e.g., for industrial or chemical processes or for diagnostic, therapeutic, or vaccination purposes.

In some embodiments of the invention, the transgenic plant host of interest is a "high expressor" of a glycoprotein described herein, including, for example, the glycoproteins comprising *N*-linked glycans that are predominately of the G0 glycan structure. By "high expressor" is intended the transgenic plant host that has been engineered to produce the glycoproteins described herein is capable of producing the glycoprotein of interest at a level such that the glycoprotein of interest represents at least 5% or more of the total soluble protein produced in the transgenic plant host. In some embodiments, a "high expressor" is a transgenic plant host that has been engineered to produce the glycoproteins described herein such that the glycoprotein of interest represents at least 5%, 6%, 7%, 8%, 9%, 10%, 11%, 12%, 13%, 14%, 15%, or more of the total soluble protein produced in the transgenic plant host. Thus, for example, in one embodiment, the transgenic plant host is a duckweed that has been modified to inhibit expression of XylT and FucT, and the transgenic duckweed is a high expressor of a glycoprotein described herein. In some of these embodiments, the transgenic duckweed is a high expressor of a glycan-optimized monoclonal antibody having the predominate G0 glycoform described herein above. In yet other embodiments, the transgenic duckweed expresses the glycan-optimized monoclonal antibody such that this glycoprotein represents about 5%, 6%, 7%, 8%, 9%, 10%, 11%, 12%, 13%, 14%, 15%, or greater, of the total soluble protein.

#### Further Humanization of Glycoproteins

In some embodiments, it may be desirable for the glycoproteins of the present invention to comprise complex *N*-glycans having the terminal  $\beta(1,4)$ -galactose

residues attached to the 1,3 and/or 1,6 mannose arms. Without being bound by theory, these terminal galactose residues may contribute to the therapeutic function and/or pharmacokinetic activity of a glycoprotein. It is recognized that the methods of the present invention can be paired with other methods known in the art to further modify the glycoproteins of the invention such that one or more of the *N*-glycans attached thereto comprises one or more terminal galactose residues, i.e., wherein one or more of the *N*-glycans is represented by the G1 or G2 glycan species. In this manner, the glycoprotein compositions of the present invention, including the monoclonal antibodies described herein, can be modified, for example, enzymatically with use of a glycosyltransferase enzyme to obtain glycoproteins having a substantially homogenous glycosylation profile for the G1, preferably for the G2 glycan species. See, for example, U.S. Patent Application Publication No. 2004/0191256, herein incorporated by reference in its entirety, teaching galactosyltransferase modification of a substrate glycoprotein to obtain a glycoprotein wherein substantially all of the *N*-linked glycan species are of the G2 form. In this manner, the glycoprotein of interest can be reacted with an activated galactose in the presence of a galactosyltransferase and a metal salt. The galactosyltransferase can be a mammalian  $\beta$ 1,4 galactosyltransferase (GalT), for example, human GalT, and the activated galactose can be, for example, UDP-galactose.

Alternatively, the transgenic plants of the invention having FucT and XylT expression silenced in the manner set forth herein can be further modified in their glycosylation machinery such that they express a galactosyltransferase and efficiently attach the terminal galactose residue to the *N*-glycans of endogenous and heterologous glycoproteins produced therein. In this manner, the transgenic plants of the invention can be further modified by introducing a nucleotide construct, for example, an expression cassette, that provides for the expression of a galactosyltransferase. The galactosyltransferase can be a mammalian  $\beta$ 1,4 galactosyltransferase (GalT), for example, human GalT (see, for example, U.S. Patent No. 6,998,267, herein incorporated by reference in its entirety) or a hybrid GalT (see, for example, WO 03/078637, herein incorporated by reference in its entirety) comprising at least a portion of a cytoplasmic tail-transmembrane-stem region of a first glycosyltransferase (e.g., a plant glycosyltransferase such as xylosyltransferase, *N*-acetylglycosaminyltransferase or fucosyltransferase) and at least a portion of a

catalytic region of a second glycosyltransferase (e.g., mammalian glycosyltransferase, for example, human GalT). By silencing expression of XylT and FucT in a plant, for example, a duckweed, and providing for expression of GalT, for example, human GalT, or a hybrid enzyme comprising a portion of the catalytic domain of GalT, for example, human GalT, in this plant, for example, duckweed, it is possible to obtain transgenic plants producing glycoproteins, both endogenous and heterologous, that have an altered glycosylation pattern, wherein the *N*-linked glycans attached thereto have a reduction in the attachment of plant-specific xylose and plant-specific fucose residues and which comprise the terminal galactose residues (i.e., G2 glycan species). In this manner, glycoproteins that have a substantially homogeneous profile for the G2 glycan species, and/or which are substantially homogeneous for the G2 glycoform can be obtained from transgenic plants of the invention.

In other embodiments, it may be desirable to further modify the glycosylation pattern of the glycoproteins of the invention, wherein the *N*-linked glycans attached thereto further comprise a terminal sialic acid residue attached to one or both of the galactose residues attached to the 1,3 and 1,6 mannose arms. The addition of the terminal sialic acid residue(s) may be required for the sustained stability, and in some cases function, of some therapeutic proteins.

Depending upon the transgenic plant system, natural sialylation of glycoproteins may occur. Thus, there have been reports in the literature that cultured *Arabidopsis*, tobacco, and *Medicago* cultured cells synthesize sialylated glycoproteins (Shah *et al.* (2003) *Nat. Biotech.* 21(12):1470-1471; Joshi and Lopez (2005) *Curr. Opin. Plant Biol.* 8(2):223-226). More recently, it was reported that Japanese rice express active sialyltransferase-like proteins (Takashima *et al.* (2006) *J. Biochem. (Tokyo)* 139(2):279-287). Hence, there are now orthogonal reports that plants have the machinery required to sialylate glycoproteins.

Where further modification of the glycosylation pattern of the glycoproteins of the invention, wherein the *N*-linked glycans attached thereto further comprise a terminal sialic acid residue attached to one or both of the galactose residues attached to the 1,3 and 1,6 mannose arms, is desired, the transgenic plants of the invention can be modified to express a  $\beta$ -1,4 galactosyltransferase, for example, human  $\beta$ -1,4 galactosyltransferase, and to express or overexpress a sialyltransferase. Thus, for example, the transgenic plants can be further modified to express a sialyltransferase

such as  $\alpha$ -2,3- and/or  $\alpha$ -2,6-sialyltransferase. See, for example WO 2004/071177; and Wee *et al.* (1998) *Plant Cell* 10:1759-1768; herein incorporated by reference in their entirety. Alternatively, the transgenic plants of the invention can be modified to express a  $\beta$ -1,4 galactosyltransferase, for example, human  $\beta$ -1,4 galactosyltransferase, and to express any other enzymes that are deficient in the plant host's sialic acid pathway. The strategy(s) employed can be determined after an initial investigation of whether the particular plant host, for example, a duckweed, naturally expresses sialic acid-containing N-glycans on native or recombinantly produced glycoproteins. For example, if there is not evidence for the presence of the terminal sialic acid residues on N-glycans of glycoproteins produced within the transgenic plant host, particularly a transgenic plant host engineered to express a  $\beta$ -1,4 galactosyltransferase, then one or both of these strategies could be employed to achieve terminal sialylation of the N-glycans of glycoproteins produced within the transgenic plant host of interest.

Alternatively, the glycoprotein compositions of the invention that are substantially homologous for G2 glycan species or the G2 glycoform can be modified by *in vitro* enzymatic processing; see, for example, U.S. Patent Application Publication No. 20030040037; herein incorporated by reference in its entirety.

It is also recognized that for some glycoproteins produced in the transgenic plants of the invention, it may be desirable to have the mammalian  $\alpha$ 1-6 fucose residue attached to the trimannose core structure ( $\text{Man}_3\text{GlcNAc}_2$ ) of the N-glycan species attached thereto. In such embodiments, the transgenic plants of the invention can be further genetically modified to express an  $\alpha$ 1-6 fucosyltransferase, for example, human  $\alpha$ 1-6 fucosyltransferase, using glycoengineering methods known in the art.

It is recognized that the glycoprotein compositions of the invention can be produced by engineering any host cell of interest, including the plant host cells exemplified and described herein. In this manner, other protein expression host systems in addition to plant hosts, including animal, insect, bacterial cells and the like may be used to produce glycoprotein compositions according to the present invention. Such protein expression host systems may be engineered or selected to express a predominant glycoform or alternatively may naturally produce glycoproteins having predominant glycan structures. Examples of engineered protein expression host systems producing a glycoprotein having a predominant glycoform include gene



knockouts/mutations (Shields *et al.* (2002) *JBC* 277:26733-26740); genetic engineering (Umana *et al.* (1999) *Nature Biotech.* 17:176-180); or a combination of both. Alternatively, certain cells naturally express a predominant glycoform, for example, chickens, humans, and cows (Raju *et al.* (2000) *Glycobiology* 10:477-486).

5 Thus, the expression of a glycoprotein, including an immunoglobulin such as a monoclonal antibody, or composition having predominantly one specific glycan structure according to the present invention can be obtained by one skilled in the art by selecting at least one of many expression host systems. Further expression host systems found in the art for production of glycoproteins include: CHO cells (see, for  
10 example, WO 9922764A1 and WO 03/035835A1); hybridoma cells (Trebak *et al.* (1999) *J. Immunol. Methods* 230:59-70); insect cells (Hsu *et al.* (1997) *JBC* 272:9062-970). See also, WO 04/074499A2 regarding additional plant host systems.

The glycoproteins produced in accordance with the methods of the present invention can be harvested from host cells in which they are recombinantly produced  
15 in order to obtain them in their isolated or purified form. In this manner, the recombinantly produced glycoproteins of the invention are isolated from the host cells using any conventional means known in the art and purified, for example, by chromatography, electrophoresis, dialysis, solvent-solvent extraction, and the like. Thus, the present invention also provides for purified glycoproteins, including  
20 monoclonal antibody compositions, where the glycoproteins have substantially homogeneous glycosylation profiles, and are substantially homogeneous for the G0 glycoform. These purified glycoproteins are substantially free of host cellular material, and include preparations of glycoprotein having less than about 30%, 20%, 10%, 5%, or 1% (by dry weight) of contaminating protein, as noted herein above. In  
25 some embodiments, these purified glycoproteins can include at least 0.001%, 0.005%, 0.1%, 0.5%, 1%, 1.5%, 2%, 2.5%, 3%, 3.5%, 4%, 4.5%, 5%, 5.5%, 6%, 6.5%, 7%, 7.5%, 8%, 8.5%, 9%, 9.5%, 10%, 15%, 20%, 25%, or up to about 30% (by dry weight) of contaminating protein. Furthermore, for the recombinantly produced purified glycoproteins of the invention, optimally culture medium represents less than  
30 about 30%, 20%, 10%, 5%, or 1% (by dry weight) of chemical precursors or non-protein-of-interest chemicals within the purified glycoprotein preparation, as noted herein above. Thus, in some embodiments, culture medium components within these purified glycoproteins can represent at least 0.001%, 0.005%, 0.1%, 0.5%, 1%, 1.5%,

2%, 2.5%, 3%, 3.5%, 4%, 4.5%, 5%, 5.5%, 6%, 6.5%, 7%, 7.5%, 8%, 8.5%, 9%, 9.5%, 10%, 15%, 20%, 25%, or up to about 30% (by dry weight) of chemical precursors or non-protein-of-interest chemicals within the purified glycoprotein preparation. In some embodiments, isolation and purification results in recovery of purified glycoprotein that is free of contaminating host protein, free of culture medium components, and/or free of both contaminating host protein and culture medium components.

Thus, in some embodiments, the protein expression host system is a plant, for example, a duckweed, and the purified glycoprotein obtained from the plant host is substantially free of plant cellular material, including embodiments where the preparations of glycoprotein have less than about 30%, 20%, 10%, 5%, or 1% (by dry weight) of contaminating plant protein. In other embodiments, the plant culture medium represents less than about 30%, 20%, 10%, 5%, or 1% (by dry weight) of chemical precursors or non-protein-of-interest chemicals within the purified glycoprotein.

In some embodiments, these purified glycoproteins obtained from the plant host can include at least 0.001%, 0.005%, 0.1%, 0.5%, 1%, 1.5%, 2%, 2.5%, 3%, 3.5%, 4%, 4.5%, 5%, 5.5%, 6%, 6.5%, 7%, 7.5%, 8%, 8.5%, 9%, 9.5%, 10%, 15%, 20%, 25%, or up to about 30% (by dry weight) of contaminating plant protein. In other embodiments, plant culture medium components within in these purified glycoproteins can represent at least 0.001%, 0.005%, 0.1%, 0.5%, 1%, 1.5%, 2%, 2.5%, 3%, 3.5%, 4%, 4.5%, 5%, 5.5%, 6%, 6.5%, 7%, 7.5%, 8%, 8.5%, 9%, 9.5%, 10%, 15%, 20%, 25%, or up to about 30% (by dry weight) of chemical precursors or non-protein-of-interest chemicals within the purified glycoprotein. In some embodiments, isolation and purification from the plant host results in recovery of purified glycoprotein that is free of contaminating plant protein, free of plant culture medium components, and/or free of both contaminating plant protein and plant culture medium components.

#### Expression Cassettes

According to the present invention, stably transformed higher plants, for example, stably transformed duckweed, are obtained by transformation with a polynucleotide of interest contained within an expression cassette. Depending upon

the objective, the polynucleotide of interest can be one encoding a FucT or XylT polypeptide of interest, for example, encoding the polypeptide set forth in SEQ ID NO:3 (FucT) or SEQ ID NO:6 or 21 (XylT), or a variant thereof, thus providing for expression of these polypeptides in a cell, for example, a plant cell, or can be a FucT or XylT inhibitory polynucleotide that is capable of inhibiting expression or function of the FucT or XylT polypeptide when stably introduced into a cell, for example, a plant cell of interest.

Thus, in some embodiments, the FucT and/or XylT polynucleotides of the invention, including those set forth in SEQ ID NOS:1 and 2 (FucT) and SEQ ID NOS:4, 5, 19, and 20 (XylT) and fragments and variants thereof, are used to construct expression cassettes that comprise a FucT and/or XylT inhibitory polynucleotide as defined herein above. Stably introducing such an expression cassette into a plant or plant cell of interest can provide for inhibition of expression or function of the FucT and/or XylT polypeptides of the invention, including those set forth in SEQ ID NO:3 (FucT) and SEQ ID NO:6 or 21 (XylT) and variants thereof, thereby altering the *N*-glycan glycosylation pattern of endogenous and heterologous glycoproteins within a plant or plant cell stably transformed with the expression cassette.

In some embodiments, the plant or plant cell that is stably transformed with an expression cassette comprising a FucT and/or XylT inhibitory polynucleotide has also been stably transformed with an expression cassette that provides for expression of a heterologous polypeptide of interest, for example, a mammalian protein of interest, including the mammalian proteins of pharmaceutical interest as noted herein above. The expression cassette providing for expression of a heterologous polypeptide of interest can be provided on the same polynucleotide (for example, on the same transformation vector) for introduction into a plant, or on a different polynucleotide (for example, on different transformation vectors) for introduction into the plant or plant cell of interest at the same time or at different times, by the same or by different methods of introduction, for example, by the same or different transformation methods.

The expression cassettes of the present invention comprise expression control elements that at least comprise a transcriptional initiation region (e.g., a promoter) linked to the polynucleotide of interest, i.e., a polynucleotide encoding a FucT or XylT polypeptide of the invention, a FucT and/or XylT inhibitory polynucleotide, or a

polynucleotide encoding a heterologous polypeptide of interest, for example, a mammalian protein. Such an expression cassette is provided with a plurality of restriction sites for insertion of the polynucleotide or polynucleotides of interest (e.g., one polynucleotide of interest, two polynucleotides of interest, etc.) to be under the transcriptional regulation of the promoter and other expression control elements. In particular embodiments of the invention, the polynucleotide to be transferred contains two or more expression cassettes, each of which encodes at least one polynucleotide of interest.

By “expression control element” is intended a regulatory region of DNA, usually comprising a TATA box, capable of directing RNA polymerase II, or in some embodiments, RNA polymerase III, to initiate RNA synthesis at the appropriate transcription initiation site for a particular coding sequence. An expression control element may additionally comprise other recognition sequences generally positioned upstream or 5' to the TATA box, which influence (e.g., enhance) the transcription initiation rate. Furthermore, an expression control element may additionally comprise sequences generally positioned downstream or 3' to the TATA box, which influence (e.g., enhance) the transcription initiation rate.

The transcriptional initiation region (e.g., a promoter) may be native or homologous or foreign or heterologous to the host, or could be the natural sequence or a synthetic sequence. By foreign, it is intended that the transcriptional initiation region is not found in the wild-type host into which the transcriptional initiation region is introduced. By “functional promoter” is intended the promoter, when operably linked to a sequence encoding a protein of interest, is capable of driving expression (i.e., transcription and translation) of the encoded protein, or, when operably linked to an inhibitory sequence encoding an inhibitory nucleotide molecule (for example, a hairpin RNA, double-stranded RNA, miRNA polynucleotide, and the like), the promoter is capable of initiating transcription of the operably linked inhibitory sequence such that the inhibitory nucleotide molecule is expressed. The promoters can be selected based on the desired outcome. Thus the expression cassettes of the invention can comprise constitutive, tissue-preferred, or other promoters for expression in plants.

As used herein a chimeric gene comprises a coding sequence operably linked to a transcription initiation region that is heterologous to the coding sequence.

Any suitable promoter known in the art can be employed according to the present invention, including bacterial, yeast, fungal, insect, mammalian, and plant promoters. For example, plant promoters, including duckweed promoters, may be used. Exemplary promoters include, but are not limited to, the Cauliflower Mosaic Virus 35S promoter, the opine synthetase promoters (e.g., nos, mas, ocs, etc.), the ubiquitin promoter, the actin promoter, the ribulose biphosphate (RubP) carboxylase small subunit promoter, and the alcohol dehydrogenase promoter. The duckweed RubP carboxylase small subunit promoter is known in the art (Silverthorne *et al.* (1990) *Plant Mol. Biol.* 15:49). Other promoters from viruses that infect plants, preferably duckweed, are also suitable including, but not limited to, promoters isolated from Dasheen mosaic virus, Chlorella virus (e.g., the Chlorella virus adenine methyltransferase promoter; Mitra *et al.* (1994) *Plant Mol. Biol.* 26:85), tomato spotted wilt virus, tobacco rattle virus, tobacco necrosis virus, tobacco ring spot virus, tomato ring spot virus, cucumber mosaic virus, peanut stunt virus, alfalfa mosaic virus, sugarcane bacilliform badnavirus and the like.

Other suitable expression control elements are disclosed in the commonly owned and copending provisional application entitled "*Expression Control Elements from the Lemnaceae Family*," assigned U.S. Patent Application No. 60/759,308, Attorney Docket No. 040989/243656, filed January 17, 2006, herein incorporated by reference in its entirety. The expression control elements disclosed in this copending application were isolated from ubiquitin genes for several members of the *Lemnaceae* family, and are thus referred to as "*Lemnaceae* ubiquitin expression control elements." SEQ ID NO:7 of the present application sets forth the full-length *Lemna minor* ubiquitin expression control element, including both the promoter plus 5' UTR (nucleotides 1-1625) and intron (nucleotides 1626-2160). SEQ ID NO:8 sets forth the full-length *Spirodella polyrrhiza* ubiquitin expression control element, including both the promoter plus 5' UTR (nucleotides 1-1041) and intron (nucleotides 1042-2021). SEQ ID NO:9 sets forth the full-length *Lemna aequinoctialis* ubiquitin expression control element, including both the promoter plus 5' UTR (nucleotides 1-964) and intron (nucleotides 965-2068). SEQ ID NO:10 sets forth the promoter plus 5' UTR portion of the *L. minor* ubiquitin expression control element (designated "LmUbq promoter" herein). SEQ ID NO:11 sets forth the promoter plus 5' UTR portion of the *S. polyrrhiza* ubiquitin expression control element (designated "SpUbq promoter")



herein). SEQ ID NO:12 sets forth the promoter plus 5' UTR portion of the *L. aequinoctialis* ubiquitin expression control element (designated "LaUbq promoter" herein). SEQ ID NO:13 sets forth the intron portion of the *L. minor* ubiquitin expression control element (designated "LmUbq intron" herein). SEQ ID NO:14 sets forth the intron portion of the *S. polyrrhiza* ubiquitin expression control element (designated "SpUbq intron" herein). SEQ ID NO:15 sets forth the intron portion of the *L. aequinoctialis* ubiquitin expression control element (designated "LaUbq intron" herein). It is recognized that the individual promoter plus 5' UTR sequences set forth in SEQ ID NOs:10-12, and biologically active variants and fragments thereof, can be used to regulate transcription of operably linked nucleotide sequences of interest in plants. Similarly, one or more of the intron sequences set forth in SEQ ID NOs:13-15, and biologically active fragments or variants thereof, can be operably linked to a promoter of interest, including a promoter set forth in SEQ ID NO:10, 11, or 12 in order to enhance expression of a nucleotide sequence that is operably linked to that promoter.

Fragments and variants of the disclosed expression control elements can also be used within expression cassettes to drive expression of the operably linked polynucleotide of interest. By "fragment of an expression control element" is intended a portion of the full-length expression control element, such as a portion of any one of the expression control elements set forth in SEQ ID NOs:7-9. Fragments of an expression control element retain biological activity and hence encompass fragments capable of initiating or enhancing expression of an operably linked polynucleotide of interest. Thus, for example, less than the entire expression control elements disclosed herein may be utilized to drive expression of an operably linked polynucleotide of interest. Specific, non-limiting examples of such fragments of an expression control element include the nucleotide sequences set forth in any one of SEQ ID NOs:10-12 (as described herein above), as well as 5' truncations of the *L. minor* ubiquitin expression control element (SEQ ID NO:7), such as nucleotides 1288-2160 of SEQ ID NO:7 (LmUbq truncated promoter No. 1) and nucleotides 1132-2160 of SEQ ID NO:1 (LmUbq truncated promoter No. 2). See the copending provisional application assigned U.S. Patent Application No. 60/759,308, herein incorporated by reference in its entirety.

The nucleotides of such fragments will usually comprise the TATA recognition sequence of the particular expression control element. Such fragments can be obtained by use of restriction enzymes to cleave the naturally occurring expression control elements disclosed herein; by synthesizing a nucleotide sequence  
5 from the naturally occurring sequence of the expression control element DNA sequence; or can be obtained through the use of polymerase chain reaction (PCR) technology. See particularly, Mullis *et al.* (1987) *Methods Enzymol.* 155:335-350, and Erlich, ed. (1989) *PCR Technology* (Stockton Press, New York).

Variants of expression control elements, such as those resulting from site-  
10 directed mutagenesis, can also be used in the expression cassettes of the present invention to provide expression of the operably linked polynucleotide of interest. By “variant of an expression control element” is intended sequences having substantial similarity with an expression control element disclosed herein (for example, the expression control element set forth in SEQ ID NO:7, 9, or 9), or with a fragment  
15 thereof (for example, the respective sequences set forth in SEQ ID NOs:10-15). Naturally occurring variants of expression control elements can be identified with the use of well-known molecular biology techniques, as, for example, with PCR and hybridization techniques as outlined above. Variant expression control elements also include synthetically derived nucleotide sequences, such as those generated, for  
20 example, by using site-directed mutagenesis. Generally, variants of a particular expression control element disclosed herein, including variants of SEQ ID NOs:7-15, will have at least 40%, 50%, 60%, 65%, 70%, generally at least 75%, 80%, 85%, preferably about 90%, 91%, 92%, 93%, 94%, to 95%, 96%, 97%, and more preferably about 98%, 99% or more sequence identity to that particular nucleotide  
25 sequence as determined by sequence alignment programs described herein above using default parameters.

Expression control elements, including promoters, can be chosen to give a desired level of regulation. For example, in some instances, it may be advantageous to use a promoter that confers constitutive expression (e.g, the mannopine synthase  
30 promoter from *Agrobacterium tumefaciens*). Alternatively, in other situations, for example, where expression of a heterologous polypeptide is concerned, it may be advantageous to use promoters that are activated in response to specific environmental stimuli (e.g., heat shock gene promoters, drought-inducible gene

promoters, pathogen-inducible gene promoters, wound-inducible gene promoters, and light/dark-inducible gene promoters) or plant growth regulators (e.g., promoters from genes induced by abscissic acid, auxins, cytokinins, and gibberellic acid). As a further alternative, promoters can be chosen that give tissue-specific expression (e.g.,  
5 root, leaf, and floral-specific promoters).

The overall strength of a given promoter can be influenced by the combination and spatial organization of cis-acting nucleotide sequences such as upstream activating sequences. For example, activating nucleotide sequences derived from the *Agrobacterium tumefaciens* octopine synthase gene can enhance transcription from  
10 the *Agrobacterium tumefaciens* mannopine synthase promoter (see U.S. Patent 5,955,646 to Gelvin *et al.*). In the present invention, the expression cassette can contain activating nucleotide sequences inserted upstream of the promoter sequence to enhance the expression of the nucleotide sequence of interest. In one embodiment, the expression cassette includes three upstream activating sequences derived from the  
15 *Agrobacterium tumefaciens* octopine synthase gene operably linked to a promoter derived from an *Agrobacterium tumefaciens* mannopine synthase gene (see U.S. Patent 5,955,646, herein incorporated by reference).

Where the expression control element will be used to drive expression of an operably linked DNA sequence encoding a small hpRNA molecule, for example,  
20 within an RNAi expression cassette described herein above, it is advantageous to use an expression control element comprising a promoter recognized by the DNA dependent RNA polymerase III. As used herein, "a promoter recognized by the DNA dependent RNA polymerase III" is a promoter which directs transcription of the associated DNA region through the polymerase action of RNA polymerase III. These  
25 include genes encoding 5S RNA, tRNA, 7SL RNA, U6 snRNA and a few other small stable RNAs, many involved in RNA processing. Most of the promoters used by Pol III require sequence elements downstream of +1, within the transcribed region. A minority of pol III templates however, lack any requirement for intragenic promoter elements. These are referred to as type 3 promoters. By "type 3 Pol III promoters" is  
30 intended those promoters that are recognized by RNA polymerase III and contain all cis-acting elements, interacting with the RNA polymerase III upstream of the region normally transcribed by RNA polymerase III. Such type 3 Pol III promoters can be

assembled within the RNAi expression cassettes of the invention to drive expression of the operably linked DNA sequence encoding the small hpRNA molecule.

Typically, type 3 Pol III promoters contain a TATA box (located between -25 and -30 in Human U6 snRNA gene) and a Proximal Sequence element (PSE; located between -47 and -66 in Human U6 snRNA). They may also contain a Distal Sequence Element (DSE; located between -214 and -244 in Human U6 snRNA). Type 3 Pol III promoters can be found, e.g., associated with the genes encoding 7SL RNA, U3 snRNA and U6 snRNA. Such sequences have been isolated from *Arabidopsis*, rice, and tomato. See, for example, SEQ ID NOs:1-8 of U.S. Patent Application Publication No. 20040231016.

Other nucleotide sequences for type 3 Pol III promoters can be found in nucleotide sequence databases under the entries for the *A. thaliana* gene AT7SL-1 for 7SL RNA (X72228), *A. thaliana* gene AT7SL-2 for 7SL RNA (X72229), *A. thaliana* gene AT7SL-3 for 7SL RNA (AJ290403), *Humulus lupulus* H17SL-1 gene (AJ236706), *Humulus lupulus* H17SL-2 gene (AJ236704), *Humulus lupulus* H17SL-3 gene (AJ236705), *Humulus lupulus* H17SL-4 gene (AJ236703), *A. thaliana* U6-1 snRNA gene (X52527), *A. thaliana* U6-26 snRNA gene (X52528), *A. thaliana* U6-29 snRNA gene (X52529), *A. thaliana* U6-1 snRNA gene (X52527), *Zea mays* U3 snRNA gene (Z29641), *Solanum tuberosum* U6 snRNA gene (Z17301; X 60506; S83742), tomato U6 small nuclear RNA gene (X51447), *A. thaliana* U3C snRNA gene (X52630), *A. thaliana* U3B snRNA gene (X52629), *Oryza sativa* U3 snRNA promoter (X79685), tomato U3 small nuclear RNA gene (X14411), *Triticum aestivum* U3 snRNA gene (X63065), and *Triticum aestivum* U6 snRNA gene (X63066).

Other type 3 Pol III promoters may be isolated from other varieties of tomato, rice or *Arabidopsis*, or from other plant species using methods well known in the art. For example, libraries of genomic clones from such plants may be isolated using U6 snRNA, U3 snRNA, or 7SL RNA coding sequences (such as the coding sequences of any of the above mentioned sequences identified by their accession number and additionally the *Vicia faba* U6snRNA coding sequence (X04788), the maize DNA for U6 snRNA (X52315), or the maize DNA for 7SL RNA (X14661)) as a probe, and the upstream sequences, preferably the about 300 to 400 bp upstream of the transcribed regions may be isolated and used as type 3 Pol III promoters. Alternatively, PCR based techniques such as inverse-PCR or TAIL™-PCR may be used to isolate the

genomic sequences including the promoter sequences adjacent to known transcribed regions. Moreover, any of the type 3 Pol III promoter sequences described herein, identified by their accession numbers and SEQ ID NOS, may be used as probes under stringent hybridization conditions or as source of information to generate PCR  
5 primers to isolate the corresponding promoter sequences from other varieties or plant species.

Although type 3 Pol III promoters have no requirement for cis-acting elements located with the transcribed region, it is clear that sequences normally located downstream of the transcription initiation site may nevertheless be included in the  
10 RNAi expression cassettes of the invention. Further, while type 3 Pol III promoters originally isolated from monocotyledonous plants can effectively be used in RNAi expression cassettes to suppress expression of a target gene in both dicotyledonous and monocotyledonous plant cells and plants, type 3 Pol III promoters originally isolated from dicotyledonous plants reportedly can only be efficiently used in  
15 dicotyledonous plant cells and plants. Moreover, the most efficient gene silencing reportedly is obtained when the RNAi expression cassette is designed to comprise a type 3 Pol III promoter derived from the same or closely related species. See, for example, U.S. Patent Application Publication No. 20040231016. Thus, where the plant of interest is a monocotyledonous plant, and small hpRNA interference is the  
20 method of choice for inhibiting expression of FucT and/or XylT, the type 3 Pol III promoter preferably is from another monocotyledonous plant, including the plant species for which the glycosylation pattern of *N*-linked glycans of a glycoprotein of interest is to be altered.

The expression cassette of the invention thus includes in the 5'-3' direction of  
25 transcription, an expression control element comprising a transcriptional and translational initiation region, a polynucleotide of interest, for example, a sequence encoding a heterologous protein of interest or a sequence encoding a FucT or XylT inhibitory sequence that, when expressed, is capable of inhibiting the expression or function of FucT and/or XylT, and a transcriptional and translational termination  
30 region functional in plants. Any suitable termination sequence known in the art may be used in accordance with the present invention. The termination region may be native with the transcriptional initiation region, may be native with the nucleotide sequence of interest, or may be derived from another source. Convenient termination



regions are available from the Ti-plasmid of *A. tumefaciens*, such as the octopine synthetase and nopaline synthetase termination regions. See also Guerineau *et al.* (1991) *Mol. Gen. Genet.* 262:141; Proudfoot (1991) *Cell* 64:671; Sanfacon *et al.* (1991) *Genes Dev.* 5:141; Mogen *et al.* (1990) *Plant Cell* 2:1261; Munroe *et al.* (1990) *Gene* 91:151; Ballas *et al.* (1989) *Nucleic Acids Res.* 17:7891; and Joshi *et al.* (1987) *Nucleic Acids Res.* 15:9627. Additional exemplary termination sequences are the pea RubP carboxylase small subunit termination sequence and the Cauliflower Mosaic Virus 35S termination sequence. Other suitable termination sequences will be apparent to those skilled in the art, including the oligo dT stretch disclosed herein above for use with type 3 Pol III promoters driving expression of a FucT and/or XlyT inhibitory polynucleotide that forms a small hpRNA structure.

Alternatively, the polynucleotide(s) of interest can be provided on any other suitable expression cassette known in the art.

Generally, the expression cassette will comprise a selectable marker gene for the selection of transformed cells or tissues. Selectable marker genes include genes encoding antibiotic resistance, such as those encoding neomycin phosphotransferase II (NEO) and hygromycin phosphotransferase (HPT), as well as genes conferring resistance to herbicidal compounds. Herbicide resistance genes generally code for a modified target protein insensitive to the herbicide or for an enzyme that degrades or detoxifies the herbicide in the plant before it can act. See DeBlock *et al.* (1987) *EMBO J.* 6:2513; DeBlock *et al.* (1989) *Plant Physiol.* 91:691; Fromm *et al.* (1990) *BioTechnology* 8:833; Gordon-Kamm *et al.* (1990) *Plant Cell* 2:603. For example, resistance to glyphosphate or sulfonylurea herbicides has been obtained using genes coding for the mutant target enzymes, 5-enolpyruvylshikimate-3-phosphate synthase (EPSPS) and acetolactate synthase (ALS). Resistance to glufosinate ammonium, boromoxynil, and 2,4-dichlorophenoxyacetate (2,4-D) have been obtained by using bacterial genes encoding phosphinothricin acetyltransferase, a nitrilase, or a 2,4-dichlorophenoxyacetate monooxygenase, which detoxify the respective herbicides.

For purposes of the present invention, selectable marker genes include, but are not limited to, genes encoding neomycin phosphotransferase II (Fraley *et al.* (1986) *CRC Critical Reviews in Plant Science* 4:1); cyanamide hydratase (Maier-Greiner *et al.* (1991) *Proc. Natl. Acad. Sci. USA* 88:4250); aspartate kinase; dihydrodipicolinate synthase (Perl *et al.* (1993) *BioTechnology* 11:715); bar gene (Toki *et al.* (1992) *Plant*

*Physiol.* 100:1503; Meagher *et al.* (1996) *Crop Sci.* 36:1367); tryptophan decarboxylase (Goddijn *et al.* (1993) *Plant Mol. Biol.* 22:907); neomycin phosphotransferase (NEO; Southern *et al.* (1982) *J. Mol. Appl. Gen.* 1:327); hygromycin phosphotransferase (HPT or HYG; Shimizu *et al.* (1986) *Mol. Cell. Biol.* 6:1074); dihydrofolate reductase (DHFR; Kwok *et al.* (1986) *Proc. Natl. Acad. Sci. USA* 83:4552); phosphinothricin acetyltransferase (DeBlock *et al.* (1987) *EMBO J.* 6:2513); 2,2-dichloropropionic acid dehalogenase (Buchanan-Wollaston *et al.* (1989) *J. Cell. Biochem.* 13D:330); acetohydroxyacid synthase (U.S. Pat. No. 4,761,373 to Anderson *et al.*; Haughn *et al.* (1988) *Mol. Gen. Genet.* 221:266); 5-enolpyruvyl-shikimate-phosphate synthase (aroA; Comai *et al.* (1985) *Nature* 317:741); haloarylnitrilase (WO 87/04181 to Stalker *et al.*); acetyl-coenzyme A carboxylase (Parker *et al.* (1990) *Plant Physiol.* 92:1220); dihydropteroate synthase (sulI; Guerineau *et al.* (1990) *Plant Mol. Biol.* 15:127); and 32 kDa photosystem II polypeptide (psbA; Hirschberg *et al.* (1983) *Science* 222:1346 (1983).

Also included are genes encoding resistance to: gentamycin (*e.g.*, aacC1, Wohlleben *et al.* (1989) *Mol. Gen. Genet.* 217:202-208); chloramphenicol (Herrera-Estrella *et al.* (1983) *EMBO J.* 2:987); methotrexate (Herrera-Estrella *et al.* (1983) *Nature* 303:209; Meijer *et al.* (1991) *Plant Mol. Biol.* 16:807); hygromycin (Waldron *et al.* (1985) *Plant Mol. Biol.* 5:103; Zhijian *et al.* (1995) *Plant Science* 108:219; Meijer *et al.* (1991) *Plant Mol. Bio.* 16:807); streptomycin (Jones *et al.* (1987) *Mol. Gen. Genet.* 210:86); spectinomycin (Bretagne-Sagnard *et al.* (1996) *Transgenic Res.* 5:131); bleomycin (Hille *et al.* (1986) *Plant Mol. Biol.* 7:171); sulfonamide (Guerineau *et al.* (1990) *Plant Mol. Bio.* 15:127); bromoxynil (Stalker *et al.* (1988) *Science* 242:419); 2,4-D (Streber *et al.* (1989) *BioTechnology* 7:811); phosphinothricin (DeBlock *et al.* (1987) *EMBO J.* 6:2513); spectinomycin (Bretagne-Sagnard and Chupeau, *Transgenic Research* 5:131).

The bar gene confers herbicide resistance to glufosinate-type herbicides, such as phosphinothricin (PPT) or bialaphos, and the like. As noted above, other selectable markers that could be used in the vector constructs include, but are not limited to, the pat gene, also for bialaphos and phosphinothricin resistance, the ALS gene for imidazolinone resistance, the HPH or HYG gene for hygromycin resistance, the EPSP synthase gene for glyphosate resistance, the HmI gene for resistance to the Hc-toxin, and other selective agents used routinely and known to one of ordinary skill in the art.

See Yarranton (1992) *Curr. Opin. Biotech.* 3:506; Chistopherson *et al.* (1992) *Proc. Natl. Acad. Sci. USA* 89:6314; Yao *et al.* (1992) *Cell* 71:63; Reznikoff (1992) *Mol. Microbiol.* 6:2419; Barkley *et al.* (1980) *The Operon* 177-220; Hu *et al.* (1987) *Cell* 48:555; Brown *et al.* (1987) *Cell* 49:603; Figge *et al.* (1988) *Cell* 52:713; Deuschle *et al.* (1989) *Proc. Natl. Acad. Sci. USA* 86:5400; Fuerst *et al.* (1989) *Proc. Natl. Acad. Sci. USA* 86:2549; Deuschle *et al.* (1990) *Science* 248:480; Labow *et al.* (1990) *Mol. Cell. Biol.* 10:3343; Zambretti *et al.* (1992) *Proc. Natl. Acad. Sci. USA* 89:3952; Baim *et al.* (1991) *Proc. Natl. Acad. Sci. USA* 88:5072; Wyborski *et al.* (1991) *Nuc. Acids Res.* 19:4647; Hillenand-Wissman (1989) *Topics in Mol. And Struc. Biol.* 10:143; Degenkolb *et al.* (1991) *Antimicrob. Agents Chemother.* 35:1591; Kleinschmidt *et al.* (1988) *Biochemistry* 27:1094; Gatz *et al.* (1992) *Plant J.* 2:397; Gossen *et al.* (1992) *Proc. Natl. Acad. Sci. USA* 89:5547; Oliva *et al.* (1992) *Antimicrob. Agents Chemother.* 36:913; Hlavka *et al.* (1985) *Handbook of Experimental Pharmacology* 78; and Gill *et al.* (1988) *Nature* 334:721. Such disclosures are herein incorporated by reference.

The above list of selectable marker genes are not meant to be limiting. Any selectable marker gene can be used in the present invention.

#### Modification of Nucleotide Sequences for Enhanced Expression in a Plant Host

Where the plant of interest is also genetically modified to express a heterologous protein of interest, for example, a transgenic plant host serving as an expression system for recombinant production of a heterologous protein, the present invention provides for the modification of the expressed polynucleotide sequence encoding the heterologous protein of interest to enhance its expression in the host plant. Thus, where appropriate, the polynucleotides may be optimized for increased expression in the transformed plant. That is, the polynucleotides can be synthesized using plant-preferred codons for improved expression. See, for example, Campbell and Gowri (1990) *Plant Physiol.* 92:1-11 for a discussion of host-preferred codon usage. Methods are available in the art for synthesizing nucleotide sequences with plant-preferred codons. See, e.g., U.S. Patent Nos. 5,380,831 and 5,436,391; Perlak *et al.* (1991) *Proc. Natl. Acad. Sci. USA* 15:3324; Iannacome *et al.* (1997) *Plant Mol. Biol.* 34:485; and Murray *et al.*, (1989) *Nucleic Acids. Res.* 17:477, herein incorporated by reference.

In some embodiments of the invention, the plant host is a member of the duckweed family, and the polynucleotide encoding the heterologous polypeptide of interest, for example, a mammalian polypeptide, is modified for enhanced expression of the encoded heterologous polypeptide. In this manner, one such modification is the synthesis of the polynucleotide encoding the heterologous polypeptide of interest using duckweed-preferred codons, where synthesis can be accomplished using any method known to one of skill in the art. The preferred codons may be determined from the codons of highest frequency in the proteins expressed in duckweed. For example, the frequency of codon usage for *Lemna gibba* is found on the web page: [http://www.kazusa.or.jp/codon/cgi-bin/showcodon.cgi?species=Lemna+gibba+\[gbpln\]](http://www.kazusa.or.jp/codon/cgi-bin/showcodon.cgi?species=Lemna+gibba+[gbpln]), and the frequency of codon usage for *Lemna minor* is found on the web page [http://www.kazusa.or.jp/codon/cgibin/showcodon.cgi?species=Lemna+minor+\[gbpln\]](http://www.kazusa.or.jp/codon/cgibin/showcodon.cgi?species=Lemna+minor+[gbpln]) and in Table 1. It is recognized that heterologous genes that have been optimized for expression in duckweed and other monocots, as well as other dicots, can be used in the methods of the invention. See, e.g., EP 0 359 472, EP 0 385 962, WO 91/16432; Perlak *et al.* (1991) *Proc. Natl. Acad. Sci. USA* 88:3324; Iannacome *et al.* (1997) *Plant Mol. Biol.* 34:485; and Murray *et al.* (1989) *Nuc. Acids Res.* 17:477, and the like, herein incorporated by reference. It is further recognized that all or any part of the polynucleotide encoding the heterologous polypeptide of interest may be optimized or synthetic. In other words, fully optimized or partially optimized sequences may also be used. For example, 40 %, 45%, 50%, 55%, 60%, 65%, 70%, 75%, 80%, 85%, 87%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, 99%, or 100% of the codons may be duckweed-preferred codons. In one embodiment, between 90 and 96 % of the codons are duckweed-preferred codons. The coding sequence of a polynucleotide sequence encoding a heterologous polypeptide of interest may comprise codons used with a frequency of at least 17% in *Lemna gibba*. In one embodiment, the modified nucleotide sequence is the human  $\alpha$ -2B-interferon encoding nucleotide sequence shown in SEQ ID NO:16, which contains 93% duckweed preferred codons.

**Table 1: *Lemna gibba*-preferred codons from GenBank Release 113**

UUU 2.2 ( 4)	UCU 0.5 ( 1)	UAU 2.2 ( 4)	UGU 0.0 ( 0)
UUC 50.5 ( 92)	UCC 31.9 ( 58)	UAC 40.1 ( 73)	UGC 17.6 ( 32)

	UUA 0.0 ( 0)	UCA 0.5 ( 1)	UAA 3.8 ( 7)	UGA 1.6 ( 3)
	UUG 2.7 ( 5)	UCG 15.4 ( 28)	UAG 0.0 ( 0)	UGG 24.2 ( 44)
5	CUU 0.5 ( 1)	CCU 6.6 ( 12)	CAU 0.5 ( 1)	CGU 1.1 ( 2)
	CUC 39.0 ( 71)	CCC 43.4 ( 79)	CAC 6.6 ( 12)	CGC 26.9 ( 49)
	CUA 1.1 ( 2)	CCA 2.2 ( 4)	CAA 4.4 ( 8)	CGA 1.1 ( 2)
	CUG 22.5 ( 41)	CCG 20.9 ( 38)	CAG 26.9 ( 49)	CGG 7.7 ( 14)
10	AUU 0.0 ( 0)	ACU 3.3 ( 6)	AAU 1.1 ( 2)	AGU 0.0 ( 0)
	AUC 33.5 ( 61)	ACC 26.4 ( 48)	AAC 37.9 ( 69)	AGC 22.0 ( 40)
	AUA 0.0 ( 0)	ACA 0.5 ( 1)	AAA 0.0 ( 0)	AGA 4.9 ( 9)
	AUG 33.5 ( 61)	ACG 9.3 ( 17)	AAG 57.1 ( 104)	AGG 6.0 ( 11)
15	GUU 9.3 ( 17)	GCU 7.1 ( 13)	GAU 1.6 ( 3)	GGU 1.1 ( 2)
	GUC 28.0 ( 51)	GCC 73.6 ( 134)	GAC 38.4 ( 70)	GGC 46.7 ( 85)
	GUA 0.0 ( 0)	GCA 5.5 ( 10)	GAA 2.2 ( 4)	GGA 1.1 ( 2)
	GUG 34.0 ( 62)	GCG 20.9 ( 38)	GAG 62.6 ( 114)	GGG 27.5 ( 50)

Other modifications can also be made to the polynucleotide encoding the heterologous polypeptide of interest to enhance its expression in a plant host of interest, including duckweed. These modifications include, but are not limited to, elimination of sequences encoding spurious polyadenylation signals, exon-intron splice site signals, transposon-like repeats, and other such well characterized sequences which may be deleterious to gene expression. The G-C content of the sequence may be adjusted to levels average for a given cellular host, as calculated by reference to known genes expressed in the host cell. When possible, the polynucleotide encoding the heterologous polypeptide of interest may be modified to avoid predicted hairpin secondary mRNA structures.

There are known differences between the optimal translation initiation context nucleotide sequences for translation initiation codons in animals and plants and the composition of these translation initiation context nucleotide sequences can influence the efficiency of translation initiation. See, for example, Lukaszewicz *et al.* (2000) *Plant Science* 154:89-98; and Joshi *et al.* (1997); *Plant Mol. Biol.* 35:993-1001. In the present invention, the translation initiation context nucleotide sequence for the translation initiation codon of the polynucleotide nucleotide of interest, for example, the polynucleotide encoding a heterologous polypeptide of interest, may be modified to enhance expression in duckweed. In one embodiment, the nucleotide sequence is modified such that the three nucleotides directly upstream of the translation initiation codon of the nucleotide sequence of interest are "ACC." In a second embodiment, these nucleotides are "ACA."



Expression of a transgene in a host plant, including duckweed, can also be enhanced by the use of 5' leader sequences. Such leader sequences can act to enhance translation. Translation leaders are known in the art and include, but are not limited to, picornavirus leaders, e.g., EMCV leader (Encephalomyocarditis 5' noncoding region; Elroy-Stein *et al.* (1989) *Proc. Natl. Acad. Sci USA* 86:6126); potyvirus leaders, e.g., TEV leader (Tobacco Etch Virus; Allison *et al.* (1986) *Virology* 154:9); human immunoglobulin heavy-chain binding protein (BiP; Macajak and Sarnow (1991) *Nature* 353:90); untranslated leader from the coat protein mRNA of alfalfa mosaic virus (AMV RNA 4; Jobling and Gehrke (1987) *Nature* 325:622); tobacco mosaic virus leader (TMV; Gallie (1989) *Molecular Biology of RNA*, 23:56); potato etch virus leader (Tomashevskaya *et al.* (1993) *J. Gen. Virol.* 74:2717-2724); Fcd-1 5' untranslated region (Dickey (1992) *EMBO J.* 11:2311-2317); RbcS 5' untranslated region (Silverthorne *et al.* (1990) *J. Plant. Mol. Biol.* 15:49-58); and maize chlorotic mottle virus leader (MCMV; Lommel *et al.* (1991) *Virology* 81:382). See also, Della-Cioppa *et al.* (1987) *Plant Physiology* 84:965. Leader sequence comprising plant intron sequence, including intron sequence from the maize alcohol dehydrogenase 1 (ADH1) gene, the castor bean catalase gene, or the Arabidopsis tryptophan pathway gene PAT1 has also been shown to increase translational efficiency in plants (Callis *et al.* (1987) *Genes Dev.* 1:1183-1200; Mascarenhas *et al.* (1990) *Plant Mol. Biol.* 15:913-920). See also copending provisional application U.S. Patent Application No. 60/759,308, wherein leader sequence comprising a duckweed intron sequence selected from the group consisting of the introns set forth in SEQ ID NOs:13-15 provides for increased translational efficiency in duckweed.

In some embodiments of the present invention, nucleotide sequence corresponding to nucleotides 1222-1775 of the maize alcohol dehydrogenase 1 gene (ADH1; GenBank Accession Number X04049), or nucleotide sequence corresponding to the intron set forth in SEQ ID NO:13, 14, or 15, is inserted upstream of the polynucleotide encoding the heterologous polypeptide of interest or the FucT and/or XylT inhibitory polynucleotide to enhance the efficiency of its translation. In another embodiment, the expression cassette contains the leader from the *Lemna gibba* ribulose-bis-phosphate carboxylase small subunit 5B gene (RbcS leader; see Buzby *et al.* (1990) *Plant Cell* 2:805-814; also see SEQ ID NO:16, 17, or 18 of the present invention).

See also, by way of example only, the expression vectors disclosed in the figures herein, wherein the RbcS leader and ADH1 intron are included as upstream regulatory sequences within an expression cassette comprising the FucT inhibitory polynucleotide (Figure 8), the XylT inhibitory polynucleotide (Figures 9 and 11), an expression cassette comprising the chimeric FucT/XylT inhibitory molecule (Figure 10), or an expression cassette comprising the coding sequence for the heterologous polypeptide, the IgG1 heavy chain of a monoclonal antibody (Figures 12, 13, and 14) or the light chain of a monoclonal antibody (Figure 14); wherein the LmUbq promoter and LmUbq intron are included as upstream regulatory sequences within an expression cassette comprising the FucT inhibitory polynucleotide (Figure 11), or an expression cassette comprising the coding sequence for the heterologous polypeptide, the IgF1 light chain of a monoclonal antibody (Figure 13); wherein the SpUbq promoter and SpUbq intron are included as upstream regulatory sequences within an expression cassette comprising the FucT inhibitory polynucleotide (Figure 13), or an expression cassette comprising the chimeric FucT/XylT inhibitory polynucleotide (Figure 12); and wherein the LaUbq promoter and LaUbq intron are included as upstream regulatory sequences in an expression cassette comprising the XylT inhibitory polynucleotide (Figure 13).

It is recognized that any of the expression-enhancing nucleotide sequence modifications described above can be used in the present invention, including any single modification or any possible combination of modifications. The phrase “modified for enhanced expression” in a plant, for example, a duckweed plant, as used herein refers to a polynucleotide sequence that contains any one or any combination of these modifications.

#### Signal Peptides

It is recognized that the heterologous polypeptide of interest may be one that is normally or advantageously expressed as a secreted protein. Secreted proteins are usually translated from precursor polypeptides that include a “signal peptide” that interacts with a receptor protein on the membrane of the endoplasmic reticulum (ER) to direct the translocation of the growing polypeptide chain across the membrane and into the endoplasmic reticulum for secretion from the cell. This signal peptide is often cleaved from the precursor polypeptide to produce a “mature” polypeptide

lacking the signal peptide. In an embodiment of the present invention, a biologically active polypeptide is expressed in the plant host of interest, for example, duckweed or other higher plant, from a polynucleotide sequence that is operably linked with a nucleotide sequence encoding a signal peptide that directs secretion of the polypeptide into the culture medium. Plant signal peptides that target protein translocation to the endoplasmic reticulum (for secretion outside of the cell) are known in the art. See, for example, U.S. Patent No. 6,020,169 to Lee *et al.* In the present invention, any plant signal peptide can be used to target the expressed polypeptide to the ER.

In some embodiments, the signal peptide is the *Arabidopsis thaliana* basic endochitinase signal peptide (amino acids 14-34 of NCBI Protein Accession No. BAA82823), the extensin signal peptide (Stiefel *et al.* (1990) *Plant Cell* 2:785-793), the rice  $\alpha$ -amylase signal peptide (amino acids 1-31 of NCBI Protein Accession No. AAA33885), or a modified rice  $\alpha$ -amylase signal sequence (SEQ ID NO:17). In another embodiment, the signal peptide corresponds to the signal peptide of a secreted duckweed protein.

Alternatively, a mammalian signal peptide can be used to target recombinant polypeptides expressed in a genetically engineered plant of the invention, for example, duckweed or other higher plant of interest, for secretion. It has been demonstrated that plant cells recognize mammalian signal peptides that target the endoplasmic reticulum, and that these signal peptides can direct the secretion of polypeptides not only through the plasma membrane but also through the plant cell wall. See U.S. Patent Nos. 5,202,422 and 5,639,947 to Hiatt *et al.* In one embodiment of the present invention, the mammalian signal peptide that targets polypeptide secretion is the human  $\alpha$ -2b-interferon signal peptide (amino acids 1-23 of NCBI Protein Accession No. AAB59402).

In one embodiment, the nucleotide sequence encoding the signal peptide is modified for enhanced expression in the plant host of interest, for example, duckweed or other higher plant, utilizing any modification or combination of modifications disclosed above for the polynucleotide sequence of interest.

The secreted biologically active polypeptide can be harvested from the culture medium by any conventional means known in the art and purified by chromatography, electrophoresis, dialysis, solvent-solvent extraction, and the like. In

this manner, purified polypeptides, as defined above, can be obtained from the culture medium.

Thus, in some embodiments, the protein expression host system is a plant, for example, a duckweed or other higher plant, and the secreted biologically active polypeptide is a glycoprotein of the invention, where the glycoprotein has a substantially homogeneous glycosylation profile, and is substantially homogeneous for the G0 glycoform. In such embodiments, any such glycoprotein that may remain within the plant material can optionally be isolated and purified as described above. The secreted glycoprotein can be obtained from the plant culture medium and purified using any conventional means in the art as noted above. In this manner, the purified glycoprotein obtained from the plant material is substantially free of plant cellular material, and includes embodiments where the preparations of glycoprotein have less than about 30%, 20%, 10%, 5%, or 1% (by dry weight) of contaminating plant protein. Where the purified glycoprotein is obtained from the plant culture medium, the plant culture medium represents less than about 30%, 20%, 10%, 5%, or 1% (by dry weight) of chemical precursors or non-protein-of-interest chemicals within the purified glycoprotein preparation.

In some embodiments, these purified glycoproteins obtained from the plant host can include at least 0.001%, 0.005%, 0.1%, 0.5%, 1%, 1.5%, 2%, 2.5%, 3%, 3.5%, 4%, 4.5%, 5%, 5.5%, 6%, 6.5%, 7%, 7.5%, 8%, 8.5%, 9%, 9.5%, 10%, 15%, 20%, 25%, or up to about 30% (by dry weight) of contaminating plant protein. In other embodiments, where the glycoprotein is collected from the plant culture medium, the plant culture medium in these purified glycoproteins can include at least 0.001%, 0.005%, 0.1%, 0.5%, 1%, 1.5%, 2%, 2.5%, 3%, 3.5%, 4%, 4.5%, 5%, 5.5%, 6%, 6.5%, 7%, 7.5%, 8%, 8.5%, 9%, 9.5%, 10%, 15%, 20%, 25%, or up to about 30% (by dry weight) of chemical precursors or non-protein-of-interest chemicals within the purified glycoprotein preparation. In some embodiments, isolation and purification from the plant host, and where secreted, from the culture medium, results in recovery of purified glycoprotein that is free of contaminating plant protein, free of plant culture medium components, and/or free of both contaminating plant protein and plant culture medium components.

Transformed Plants and Transformed Duckweed Plants and Duckweed Nodule Cultures

Transformation protocols as well as protocols for introducing nucleotide sequences into plants may vary depending on the type of plant or plant cell or nodule, that is, monocot or dicot, targeted for transformation. Suitable methods of introducing nucleotide sequences into plants or plant cells or nodules include microinjection (Crossway *et al.* (1986) *Biotechniques* 4:320-334), electroporation (Riggs *et al.* (1986) *Proc. Natl. Acad. Sci. USA* 83:5602-5606), *Agrobacterium*-mediated transformation (U.S. Patent Nos. 5,563,055 and 5,981,840, both of which are herein incorporated by reference), direct gene transfer (Paszkowski *et al.* (1984) *EMBO J.* 3:2717-2722), ballistic particle acceleration (see, *e.g.*, U.S. Patent Nos. 4,945,050; 5,879,918; 5,886,244; and 5,932,782 (each of which is herein incorporated by reference); and Tomes *et al.* (1995) "Direct DNA Transfer into Intact Plant Cells via Microprojectile Bombardment," in *Plant Cell, Tissue, and Organ Culture: Fundamental Methods*, ed. Gamborg and Phillips (Springer-Verlag, Berlin); McCabe *et al.* (1988) *Biotechnology* 6:923-926). The cells that have been transformed may be grown into plants in accordance with conventional ways. See, for example, McCormick *et al.* (1986) *Plant Cell Reports* 5:81-84.

The stably transformed duckweed utilized in this invention can be obtained by any method known in the art. In one embodiment, the stably transformed duckweed is obtained by one of the gene transfer methods disclosed in U.S. Patent No. 6,040,498 to Stomp *et al.*, herein incorporated by reference. These methods include gene transfer by ballistic bombardment with microprojectiles coated with a nucleic acid comprising the nucleotide sequence of interest, gene transfer by electroporation, and gene transfer mediated by *Agrobacterium* comprising a vector comprising the nucleotide sequence of interest. In one embodiment, the stably transformed duckweed is obtained via any one of the *Agrobacterium*-mediated methods disclosed in U.S. Patent No. 6,040,498 to Stomp *et al.* The *Agrobacterium* used is *Agrobacterium tumefaciens* or *Agrobacterium rhizogenes*.

It is preferred that the stably transformed duckweed plants utilized in these methods exhibit normal morphology and are fertile by sexual reproduction. Preferably, transformed plants of the present invention contain a single copy of the transferred nucleic acid, and the transferred nucleic acid has no notable



rearrangements therein. Also preferred are duckweed plants in which the transferred nucleic acid is present in low copy numbers (i.e., no more than five copies, alternately, no more than three copies, as a further alternative, fewer than three copies of the nucleic acid per transformed cell).

5           The following examples are offered by way of illustration and not by way of limitation.

## EXPERIMENTAL

10           *Lemna*, a small aquatic plant, is a scalable and economically attractive expression platform for the manufacture of therapeutic proteins free of human pathogens and with a clear path towards regulatory approval. The *Lemna* expression system (LEX System<sup>SM</sup>) enables rapid clonal expansion of transgenic plants, secretion of transgenic proteins, high protein yields, ease of containment that is comparable to mammalian cell culture systems such as CHO cells, and has the additional advantage  
15 of low operating and capital costs (Gasdaska *et al.* (2003) *Bioprocessing J.* 50-56). In addition, this plant expression system offers the advantage of high protein yields (in the range of 6-8% of the total soluble protein (TSP)). These expression levels, in combination with *Lemna*'s high protein content and fast growth rate (36 hr doubling time), enable production of >1g of mAb per kg biomass in a robust and well-  
20 controlled format.

          The following examples demonstrate how humanization of the glycosylation profile of a mAb was accomplished by coexpression of the mAb with an interference RNA (RNAi) construct targeting the endogenous expression of  $\alpha$ 1,3-fucosyltransferase and  $\beta$ 1,2-xylosyltransferase genes. The resultant mAb contained a  
25 single major N-glycan species (>95%) devoid of the plant specific  $\alpha$ -1,3-linked fucose and  $\beta$ -1,2-linked xylose sugars. In receptor binding assays, this glycan optimized mAb exhibited enhanced effector cell receptor binding activity when compared to mAb produced in wild-type *Lemna* having the native glycosylation machinery and mAb produced in CHO cells.

### 30           Example 1: Isolation of *Lemna minor* Proteins Involved in N-Glycosylation of Proteins

          In order to generate recombinant proteins with remodeled N-glycan, alpha 1-3  
35 fucosyltransferase and  $\beta$ 1-2 xylosyltransferase were selected as targets for RNAi gene

silencing in *L. minor*. Initial results from cDNA sequencing efforts indicated that two or more isoforms were present for each of the target genes. Sequence homology between the isoforms was determined to be between 90% and 95%. Full length cDNA sequences for both target genes were retrieved and characterized. The full-length cDNA sequence, including 5'- and 3'-UTR, for *L. minor*  $\alpha$ 1-3 fucosyltransferase (FucT) is set forth in Figure 1; see also SEQ ID NO:1 (open reading frame set forth in SEQ ID NO:2). The predicted amino acid sequence encoded thereby is set forth in SEQ ID NO:3. The encoded protein shares some similarity with other FucTs from other higher plants. See Figure 2. For example, the *L. minor* FucT sequence shares approximately 50.1% sequence identity with the *Arabidopsis thaliana* FucT shown in Figure 2.

The full-length cDNA sequence, including 5'- and 3'-UTR, for *L. minor*  $\beta$ 1-2 xylosyltransferase (XylT) (isoform #1) is set forth in Figure 3; see also SEQ ID NO:4 (ORF set forth in SEQ ID NO:5). The predicted amino acid sequence encoded thereby is set forth in SEQ ID NO:6. The encoded protein shares some similarity with other XylTs from other higher plants. See Figure 4. For example, the *L. minor* XylT shares approximately 56.4% sequence identity with the *Arabidopsis thaliana* XylT shown in Figure 4. A partial-length cDNA sequence, including 3'-UTR, for *L. minor*  $\beta$ 1-2 xylosyltransferase (XylT) (isoform #2) is set forth in Figure 31; see also SEQ ID NO:19 (ORF set forth in SEQ ID NO:20). The predicted amino acid sequence encoded thereby is set forth in SEQ ID NO:21. The partial-length XylT isoform #2 shares high sequence identity with the corresponding region of the full-length XylT isoform #1, as can be seen from the alignment shown in Figure 32.

#### Example 2: RNAi Inhibition of Expression of *L. minor* FucT and XylT

Several RNAi strategies were undertaken to inhibit expression of the *L. minor* FucT and XylT isoforms. Figures 5-7, 33, and 34 outline these strategies. Figures 8-13 show maps of the various constructs that were made to achieve the desired knockout of expression of these two genes. A number of transgenic lines comprising the various knockout RNAi constructs were generated using standard transformation protocols described herein above.

The test antibody, designated herein as mAbI, was expressed in wild-type *Lemna* having the native glycosylation machinery, and transgenic *Lemna* lines

expressing RNAi constructs designed to inhibit expression of *L. minor* XylT and FucT isoforms. Generally, three binary vectors were constructed for expression of mAbI in the *Lemna* system. Expression vector mAbI01 contained codon optimized genes encoding heavy (H) and light (L) chains of mAbI; vector mAbI04 contained  
 5 codon optimized genes encoding mAbI H and L chains and a chimeric RNAi construct targeting expression of both XylT and FucT isoforms; and vector mAbI05 contained codon optimized genes encoding mAbI H and L chains, a single-gene RNAi construct targeting FucT gene expression, and a single-gene RNAi construct targeting XylT gene expression. Independent transgenic lines were generated for the  
 10 mAbI01, mAbI04, and mAbI05 expression vectors.

Optimized genes for mAbI H and L chains were designed to have *Lemna* preferred codon usage (63% – 67% GC content) and contain the rice  $\alpha$ -amylase signal sequence (GenBank M24286) fused to the 5' end of their coding sequences. Restriction endonuclease sites were added for cloning into *Agrobacterium* binary  
 15 vectors (EcoRI (5') / SacI (3'), H-chain) and (Sall (5') / HindIII (3'), L-chain).

For the XF02 data presented in Figures 15-17 and the mAbI04 data presented in Figures 22-24 and 26, described herein below, the RNAi strategy for inhibiting expression of the *L. minor* FucT and XylT isoforms employed the chimeric RNAi  
 20 design shown in Figure 34. For the mAbI05 data presented in Figure 27, described herein below, the RNAi strategy for inhibiting expression of the *L. minor* FucT and XylT isoforms employed a double knockout of these genes using a combination of the single gene RNAi designs shown in Figure 5 (FucT RNAi design) and Figure 33 (XylT RNAi design).

Independent expression cassettes containing promoter, gene of interest, and Nos terminator were created for the optimized mAbI H and L chains and the single-gene or chimeric RNAi. Expression cassettes were cloned into a modification of the *Agrobacterium* binary vector pBMSP3 (obtained from Dr. Stan Gelvin, Purdue  
 25 University) with the appropriate restriction sites. Depending upon the expression cassette, the L chain was fused to either the modified chimeric octopine and mannopine synthase promoter with *Lemna gibba* 5' RbcS leader (mAbI01, Figure 14) or the high expression, constitutive *Lemna minor* polyubiquitin promoter (LmUbq) (mAbI04, Figure 12; mAbI05; Figure 13). The H-chain was fused to the modified  
 30

chimeric octopine and mannopine synthase promoter with *Lemna gibba* 5' RbcS leader (mAbI04, mAbI05, and mAbI01). The chimeric RNAi cassette, taken from plasmid XF02 in T7-4, was fused to the high expression, constitutive *Spirodela polyrhiza* polyubiquitin promoter (SpUbq). The single-gene RNAi cassette for expression of the FucT inhibitory sequence was driven by the SpUbq promoter; and the single gene RNAi cassette for expression of the XylT inhibitory sequence was driven by an operably linked expression control element comprising the *Lemna aequinoctialis* ubiquitin promoter plus 5' UTR (LaUbq promoter). The H, L, and chimeric RNAi expression cassettes were cloned into the modified pBMSP3 binary vector in tandem orientation creating plasmid mAbI04. The H, L, and single-gene RNAi expression cassettes targeting FucT and XylT expression were cloned into the modified pBMSP3 binary vector creating plasmid mAbI05. The H and L expression cassettes were cloned into the modified pBMSP3 binary vector creating plasmid mAbI01.

Though any transformation protocol can be used as noted herein above, in some embodiments, the transformation protocol was as follows. Using *Agrobacterium tumefaciens* C58Z707, a disarmed, broad host range C58 strain (Hepburn *et al.* (1985) *J. Gen. Microbiol.* 131:2961-2969), transgenic plants representing individual clonal lines were generated from rapidly growing *Lemna* minor nodules according to the procedure of Yamamoto *et al.* (2001) *In Vitro Cell Dev. Biol. Plant* 37:349-353. For transgenic screening, individual clonal lines were preconditioned for 1 week at 150 to 200  $\mu\text{mol m}^{-2}\text{s}^{-2}$  in vented plant growth vessels containing SH media (Schenk and Hildebrandt (1972) *Can. J. Botany* 50:199-204) without sucrose. Fifteen to twenty preconditioned fronds were then placed into vented containers containing fresh SH media, and allowed to grow for two weeks. Tissue and media samples from each line were frozen and stored at  $-70^{\circ}\text{C}$ .

A MALDI-TOF assay was developed to measure *L. minor*  $\beta$ -1,2-xylosyltransferase (XylT) and  $\alpha$ -1,3-fucosyltransferase (FucT) activities (Example 3 below).

Figures 15-17 represent primary screening data for the XF02, mAbII04, and mAbI05 plants lines using the aforementioned assay. In this assay, WT (wild-type) represents the FucT and XylT activity in wild-type plants while BWT (boiled wild-type) represents their activity in boiled plant extracts. Boiled wild-type (BWT) plant

extracts are representative of plant material in which FucT and XylT activity has been deactivated. This data set shows that several plant lines from each construct have a reduced level of FucT and XylT activity compared with wild-type plant lines (WT) and a comparable level of activity with boiled wild-type samples (BWT).

5 Specifically, primary screening data for transgenic RNAi *L. minor* plant lines comprising the XF02 construct of Figure 10 are shown in Figures 15 and 16. The XF02 construct expresses a chimeric RNAi molecule that targets expression of both the *L. minor* FucT and XylT proteins, including the various isoforms of the respective proteins.

10 Figure 17 shows primary screening data for transgenic RNAi *L. minor* plant lines comprising the mAbI04 construct of Figure 12 and mAbI05 construct of Figure 13.

15 Example 3: MALDI-TOF Assay for *N*-glycan  $\beta$ -1,2-xylosyltransferase (XylT) and  $\alpha$ -1,3-fucosyltransferase (FucT) Activity

The following modified MALDI-TOF assay was used to determine XylT and FucT activity in the transgenic plants described in Example 2 above.

20 **Materials**

HOMOGENIZATION BUFFER: 50 mM HEPES, pH 7.5, 0.25 M sucrose, 2 mM EDTA, 1 mM DTT.

25 REACTION BUFFER: 0.1 M Mes, pH 7.0, 10 mM MnCl<sub>2</sub>, 0.1% (v/v) Triton X-100.

URIDINE-5'-DIPHOSPHO-D-XYLOSE (UDP-Xyl)

30 GUANOSINE-5'-DIPHOSPHO-L-FUCOSE (GDP-Fuc)

*N*-ACETYLGLUCOSAMINE

POLYETHYLENE GLYCOL (PEG) MIXTURE 1000-3000 (10 mg/mL PEG 1000, 2000, and 3000 (4:5:6 ratio) mixed 4:1 with 2 mg/mL sodium iodide).

35 [Glu<sup>1</sup>]-FIBRINOPEPTIDE B (GFP), HUMAN (1 pmol/ $\mu$ L in water)

DABSYLATED, TETRAPEPTIDE, *N*-GLYCAN ACCEPTOR (EMD Biosciences)



CHCA ( $\alpha$ -CYANO-4-HYDROXYCINNAMIC ACID) MATRIX (10 mg in 50%  
[v/v] acetonitrile, 0.05% [v/v] trifluoroacetic  
acid).

5

#### Microsome preparation

*L. minor* tissue (100 mg) was ground in 1 mL of cold homogenization buffer  
in a bead mill at 5x speed for 40 s. The homogenate was spun at 1,000g for 5 min,  
4°C. The supernatant was removed and spun at 18,000g for 2 h, 4°C. The  
10 supernatant was then discarded. The pellet was resuspended in 20  $\mu$ L of cold reaction  
buffer and kept on ice or stored at -80°C until use.

#### Reaction conditions

15 The reaction mix contains 125 mM *N*-acetylglucosamine, 6.25 mM UDP-Xyl,  
6.25 mM GDP-Fuc, 12.5 mM MnCl<sub>2</sub>, and 1.5 nmol of dabsylated, tetrapeptide *N*-  
glycan acceptor. Microsomes (4  $\mu$ L) were added to the reaction mix to start the  
reaction. The reaction was incubated for 30 min at room temperature, and 90 min at  
37°C. The reaction was terminated by centrifugation at 18,000g for 1 min and  
20 incubation at 4°C.

#### MALDI-TOF analysis

A portion of the supernatant from each reaction (0.5  $\mu$ L) was mixed with 0.5  
 $\mu$ L of CHCA matrix on a MALDI target plate and allowed to dry. The MALDI  
25 instrument was set to reflectron positive ion mode and calibrated with PEG 1000-  
3000. Combined MS spectra (~200 shots) were taken from 1500-2500 Da using 0.5  
pmol GFP as the lock mass. Ion counts of the reference peak ( $m/z$  = 2222.865)  
should be above 400. Ion counts of the XylT and FucT products ( $m/z$  = 2192.854 and  
2206.870, respectively) were normalized to the reference peak and the protein  
30 concentration of the microsome fraction.

#### Example 4: Effect of RNAi Inhibition of Expression of *L. minor* FucT and XylT on Glycosylation Profile of Monoclonal Antibodies

35 Monoclonal antibodies produced by wild-type (i.e., FucT and XylT expression  
not silenced) *L. minor* comprising the mAbI01 construct (see Figure 14) and *L. minor*

lines transgenic for the mAbI04 construct (see Figure 12) or mAbI05 construct (see Figure 13) were analyzed for their *N*-glycosylation profile. The following procedures were used.

5      *Purification of mAb from Lemna.*

Plant tissue was homogenized with 50mM Sodium Phosphate, 0.3M Sodium Chloride, and 10mM EDTA at pH 7.2 using a Silverson High Shear Mixer at a tissue: buffer ratio of 1:8. The homogenate was acidified to pH 4.5 with 1M Citric Acid, and centrifuged at 7,500 x g for 30 minutes at 4°C. The supernatant was filtered through a  
10      0.22µm filter and loaded directly on mAbSelect SuRe resin (GE Healthcare) equilibrated with a solution containing 50mM Sodium Phosphate, 0.3M Sodium Chloride, and 10mM EDTA, pH 7.2. After loading, the column was washed to baseline with the equilibration buffer followed by an intermediate wash with 5 column volumes of 0.1M Sodium Acetate, pH 5.0, and finally, bound antibody was  
15      eluted with 10 column volumes of 0.1M Sodium Acetate, pH 3.0. The eluate was immediately neutralized with 2M Tris base.

*Purification of N-linked glycans.*

Protein A-purified monoclonal antibodies (1 mg) from wild-type and RNAi *L. minor* plant lines were dialyzed extensively against water and lyophilized to dryness. Samples were resuspended in 100 µL of 5% (v/v) formic acid, brought to 0.05 mg/mL pepsin, and incubated at 37°C overnight. The samples were heat inactivated at 95°C for 10 min and dried. Pepsin digests were resuspended in 100 µL of 100 mM sodium acetate, pH 5.0 and incubated with 1 mU of N-glycosidase A at 37°C overnight. The  
25      released *N*-glycans were isolated using 4 cc Carbograph SPE columns according to Packer *et al.* (1998) *Glycoconj. J.* 15: 737-747, and dried.

Dried *N*-glycans were further purified using 1 cc Waters Oasis MCX cartridges. Columns were prepared by washing with 3 column volumes of methanol followed by 3 column volumes of 5% (v/v) formic acid. *N*-glycans, resuspended in 1  
30      mL of 5% (v/v) formic acid, were loaded onto the prepared columns. The unbound fraction as well as 2 additional column volume washes of 5% (v/v) formic acid were collected, pooled and dried.

*Derivatization of oligosaccharides with 2-aminobenzoic acid (2-AA).*

Purified *N*-glycans or maltooligosaccharides were labeled with 2-AA and purified using 1 cc Waters Oasis HLB cartridges according to Anumula and Dhume (1998) *Glycobiology* 8: 685-694. Labeled *N*-glycans and maltooligosaccharides were resuspended in 50  $\mu$ L of water and analyzed by MALDI-TOF MS and normal phase (NP) HPLC-QTOF MS.

*MALDI-TOF Mass Spectrometry.*

MALDI-TOF MS was conducted using a Waters MALDI Micro MX (Millford, MA). 2-AA labeled *N*-glycans (0.5  $\mu$ L) were properly diluted with water, mixed with 0.5  $\mu$ L of 10 mg/mL DHB matrix in 70% (v/v) acetonitrile, spotted onto a target plate and analyzed in negative reflectron mode.

*NP-HPLC-Q-TOF MS analysis of 2-AA labeled N-glycans.*

2-AA labeled *N*-glycans or maltooligosaccharides were brought to 80% (v/v) acetonitrile and separated on a Waters 2695 HPLC system fitted with a TSK-Gel Amide-80 (2 mm x 25 cm, 5  $\mu$ m) column (Tosoh Biosciences, Montgomeryville, PA). 2-AA labeled carbohydrates were detected and analyzed by fluorescence (230 nm excitation, 425 nm emission) using a Waters 2475 fluorescence detector and a Waters Q-TOF API US quadropole-time of flight (Q-TOF) mass spectrometer (Millford, MA) fitted in-line with the HPLC system.

Separations were conducted at 0.2 mL/min, 40°C, using 10 mM ammonium acetate, pH 7.3 (solvent A) and 10 mM ammonium acetate, pH 7.3, 80% (v/v) acetonitrile (solvent B). Sample elution was carried out at 0% A isocratic for 5 min, followed by a linear increase to 10% A at 8 min, and a linear increase to 30% A at 48 min. The column was washed with 100% A for 15 min and equilibrated at 0% A for 15 min prior to the next injection.

Q-TOF analysis was conducted in negative ion mode with source and desolvation temperatures of 100°C and 300°C, respectively, and capillary and cone voltages of 2,100 and 30 V, respectively. Mass spectra shown are the result of combining  $\geq 50$  individual scans per labeled *N*-glycan.

*RP-HPLC-Q-TOF MS analysis of intact IgG.*

Protein A purified IgG's (50 µg) were desalted using the Waters 2695 HPLC system fitted with a Poros R1-10 column (2 mm x 30 mm; Applied Biosystems). IgG's were detected and analyzed using a Waters 2487 dual wavelength UV detector (280 nm) and the Waters Q-TOF API US. Separations were conducted at 0.15 mL/min, 60°C, using 0.05% (v/v) trifluoroacetic acid (TFA; solvent A) and 0.05% (v/v)TFA, 80% (v/v) acetonitrile (solvent B). Sample elution was carried out using a linear increase from 30 to 50% B for 5 min, an increase to 80% B for 5 min. The solvent ratio remained at 80% B for an additional 4 min, followed by a wash with 100% B for 1 min and equilibration of the column with 30% B for 15 min prior to the next run.

Q-TOF analysis was conducted in positive ion mode with source and desolvation temperatures of 100°C and 300°C, respectively, and capillary and cone voltages of 3.0 and 60 V, respectively. Data are the result of combining  $\geq 100$  individual scans and deconvolution to the parent mass spectrum using MaxEnt 1.

See also Triguero *et al.* (2005) *Plant Biotechnol. J.* 3: 449-457; Takahashi *et al.* (1998) *Anal. Biochem.* 255: 183-187; Dillon *et al.* (2004) *J. Chromatogr. A.* 1053: 299-305.

*Results.*

Figure 18 shows the structure and molecular weight of derivatized wild-type *L. minor* monoclonal antibody N-glycans.

Figure 19 shows that the wild-type mAbI01 construct (shown in Figure 15) providing for expression of the mAbI monoclonal IgG1 antibody in *L. minor*, without RNAi suppression of *L. minor* FucT and XylT, produces an N-glycosylation profile with three major N-glycan species, including one species having the  $\beta$ 1,2-linked xylose and one species having both the  $\beta$ 1,2-linked xylose and core  $\alpha$ 1,3-linked fucose residues; this profile is confirmed with liquid chromatography mass spectrometry (LC-MS) (Figure 20) and MALDI (Figure 21) analysis.

Figure 22 shows an overlay of the relative amounts of the various N-glycan species of the mAbI produced in the wild-type *L. minor* line comprising the mAbI01 construct (no suppression of FucT or XylT) and in the two transgenic *L. minor* lines comprising the mAbI04 construct of Figure 12. Note the enrichment of the GnGn

(i.e., G0) glycan species, with no  $\beta$ 1,2-linked xylose or core  $\alpha$ 1,3-linked fucose residues attached, and the absence of the species having the  $\beta$ 1,2-linked xylose or both the  $\beta$ 1,2-linked xylose and core  $\alpha$ 1,3-linked fucose residues. This profile is confirmed with mass spec (LC-MS) (Figure 23) and MALDI (Figure 24) analysis.

5           Figure 25 shows intact mass analysis of the mAbI compositions produced in wild-type *L. minor* (line 20) comprising the mAbI01 construct. When XylT and FucT expression are not suppressed in *L. minor*, the recombinantly produced mAbI composition is heterogeneous, comprising at least 9 different glycoforms, with the G0XF<sup>3</sup> glycoform being the predominate species present. Note the very minor peak  
10           representing the G0 glycoform.

          Figure 26 shows intact mass analysis of the mAbI compositions produced in transgenic *L. minor* (line 15) comprising the mAbI04 construct of Figure 12. When XylT and FucT expression are suppressed in *L. minor* using this chimeric RNAi construct, the intact mAbI composition is substantially homogeneous for G0 *N*-  
15           glycans, with only trace amounts of precursor *N*-glycans present (represented by the GnM and MGn precursor glycan species). In addition, the mAbI composition is substantially homogeneous for the G0 glycoform, wherein both glycosylation sites are occupied by the G0 *N*-glycan species, with three minor peaks reflecting trace amounts of precursor glycoforms (one peak showing mAbI having an Fc region wherein the  
20           C<sub>H</sub>2 domain of one heavy chain has a G0 glycan species attached to Asn 297, and the C<sub>H</sub>2 domain of the other heavy chain is unglycosylated; another peak showing mAbI having an Fc region wherein the C<sub>H</sub>2 domain of one heavy chain has a G0 glycan species attached to Asn 297, and the C<sub>H</sub>2 domain of the other heavy chain has the GnM or MGn precursor glycan attached to Asn 297; and another peak showing mAbI  
25           having an Fc region wherein the Asn 297 glycosylation site on each of the C<sub>H</sub>2 domains has a G0 glycan species attached, with a third G0 glycan species attached to an additional glycosylation site within the mAbI structure).

          Figure 27 shows intact mass analysis of the mAbI compositions produced in transgenic *L. minor* (line 72) comprising the mAbI05 construct of Figure 13. When  
30           XylT and FucT expression are suppressed in *L. minor* using this construct, the intact mAbI composition is substantially homogeneous for G0 *N*-glycans, with only trace amounts of precursor *N*-glycan species present (represented by the GnM and MGn precursor glycan species). In addition, the mAbI composition is substantially



homogeneous (at least 90%) for the G0 glycoform, with the same three minor peaks reflecting precursor glycoforms as obtained with the mAbI04 construct.

The receptor binding activity of the mAbI produced in the wild-type *Lemna* lines comprising the mAbI01 construct (i.e., without inhibition of XylT and FucT expression) and the transgenic *Lemna* lines comprising the mAbI04 or mAbI05 construct (i.e., with XylT and FucT expression inhibited) was compared to the receptor binding activity of the mAbI produced in mammalian cell lines (CHO and SP2/0).

Binding to Fc $\gamma$ RIIIa on freshly isolated human NK cells was assessed for the various mAbI products. Control data collected for CHO-derived mAbI and SP2/0-derived mAbI is shown in Figure 35. Test data collected for wild-type *Lemna*-produced mAbI having the normal plant *N*-glycan profile are designated as mAbI01-15 and mAbI01-20, wherein the mAbI product has *N*-linked glycans that include  $\alpha$ (1,3)-fucose residues (see Figures 36 and 37). Test data collected for transgenic *Lemna*-derived mAbI having an optimized *N*-glycan profile (OPT) obtained with gene-silencing RNAi constructs that target expression of  $\alpha$ 1,3-fucosyltransferase are designated as mAbI05-72, mAbI05-74, mAbI04-24, and mAbI04-15 (see Figures 36 and 37), wherein the mAbI product has *N*-linked glycans that are devoid of  $\alpha$ (1,3)-fucose residues. Data comparing binding efficacy of mAbI01-15, mAbI01-20, mAbI SP2/0, mAbI04-15, mAbI04-24, mAbI05-72, and mAbI05-74 to recombinant mouse Fc $\gamma$ RIV, a receptor that is sensitive to IgG fucose levels and which served as a surrogate for human Fc $\gamma$ RIIIa, is shown in Figure 38.

These data demonstrate that the transgenic *Lemna*-derived mAbI product having the optimized glycan profile (OPT) shows enhanced binding to Fc $\gamma$ RIIIa on freshly isolated human NK cells (enhanced about 20 to 50-fold) as well as enhanced binding to recombinant mouse Fc $\gamma$ RIV (enhanced about 10-fold) as compared to the wild-type *Lemna*-derived mAbI product.

#### Example 6: Production of Anti-CD30 Monoclonal Antibody Having Improved Receptor Binding and Increased ADCC Activity

This example outlines the expression of human anti-CD30 mAbs in *Lemna*. Optimization of anti-CD30 mAb glycosylation was accomplished by co-expression with an RNAi construct targeting the endogenous expression of  $\alpha$ -1,3-

fucosyltransferase (FucT) and  $\beta$ -1,2-xylosyltransferase (XylT) genes in a manner similar to that noted in the examples above for mAbI. The resultant anti-CD30 mAb produced in *Lemna* having its native glycosylation machinery engineered to suppress FucT and XylT expression contained a single major *N*-glycan species without any trace of plant-specific *N*-glycans. In addition to the *N*-glycan homogeneity, glyco-optimized anti-CD30 mAbs were also shown to have enhanced antibody-dependent cell-mediated cytotoxicity (ADCC) and effector cell receptor binding activity when compared to CHO-expressed anti-CD30 mAbs.

## METHODS

### *Strains and Reagents.*

Novablue competent *Escherichia coli* cells were used for all recombinant DNA work (EMD Biosciences, San Diego, CA). Restriction endonucleases and DNA modification enzymes were obtained from New England Biolabs (Ipswich, MA).

Oligonucleotides were obtained from Integrated DNA technologies (Coralville, IA). Waters Oasis HLB and MCX columns (1 cc), 2,5-dihydroxybenzoic acid (DHB), and  $\alpha$ -cyano-4-hydroxycinnamic acid (CHCA) were from Waters Corporation (Milford, MA). Purified dabsylated, tetrapeptide, GnGn *N*-glycan acceptors (GnGn-dabsyl-peptide) and *N*-glycosidase A were from EMD Biosciences. Carbograph SPE columns (4 cc) were from Grace Davidson Discovery Sciences (Deerfield, IL). Uridine-5'-diphospho-D-xylose (UDP-Xyl) was purchased from Carbosource Services (Athens, GA). Acetonitrile (Optima grade) was from Fisher Scientific (Summerville, NY). Ammonium acetate was from MP Biochemicals (Irvine, CA). Maltooligosaccharides (MD6-1) were from V-Labs Inc. (Covington, CA).

Monosaccharide standards were from Dionex (Sunnyvale, CA). BATDA (bis(acetoxymethyl)2,2':6',2"-terpyridine-6,6"-dicarboxylate) and Europium solution were from Perkin-Elmer (Wellesley, MA). Guanosine-5'-diphospho-L-fucose (GDP-Fuc), *N*-acetylglucosamine (GlcNAc), 2-aminobenzoic acid (2-AA) and all other materials were from Sigma (St. Louis, MO).

### *Construction of mAb and RNAi expression vectors.*

The heavy (H) and light (L) chain variable region cDNA sequences of fully human mAb1 kappa antibody MDX-060 derived from a transgenic Medarex HuMAb-

Mouse® (Borchmann *et al.* (2003) *Blood* 102:3737-3742) were determined and the full length MDX-060 human mAb1 kappa antibody was produced recombinantly by a Chinese hamster ovary cell line, CHO DG44 (Urlaub *et al.* (1986) *Cell Mol. Genet.* 12:555-566), using standard techniques. Optimized genes for H and L chains were designed to have *Lemna*-preferred codon usage (63%-67% GC content) and contain the rice  $\alpha$ -amylase signal sequence (GenBank M24286) fused to the 5' end of their coding sequences. Restriction endonuclease sites were added for cloning into *Agrobacterium* binary vectors (EcoRI (5') / SacI (3'), H-chain) and (SalI (5') / HindIII (3'), L-chain). Synthetic genes were constructed and provided by Picoscript (Houston, TX).

A chimeric hairpin RNA (see Figure 34) was designed to target silencing of endogenous *Lemna* genes encoding  $\alpha$ -1,3-fucosyltransferase (based on the coding sequence for *L. minor* FucT isoform #1, set forth in SEQ ID NO:2; see also GenBank DQ789145) and  $\beta$ -1,2-xylosyltransferase (based on the coding sequence for *L. minor* XylT isoform #2, nt 1-1275 of SEQ ID NO:19; set forth in SEQ ID NO:20; see also GenBank DQ789146). The chimeric FucT + XylT hairpin RNA was designed to have 602 bp of double stranded FucT sequence, 626 bp of double stranded XylT sequence, and 500 bp of spacer sequence. The sense strand portion of the hairpin RNA cassette encompasses the FucT forward fragment (nt 12-613 of SEQ ID NO:2; equivalent to nt 254-855 of SEQ ID NO:1) and XylT forward fragment (nt 1-626 of SEQ ID NO:19), a spacer sequence (nt 627-1126 of SEQ ID NO:19). The antisense strand portion of the hairpin RNA encompasses the XylT reverse fragment (antisense version of nt 1-626 SEQ ID NO:19) and FucT reverse fragment (antisense version of nt 12-613 of SEQ ID NO:2 or nt 254-855 of SEQ ID NO:1). The chimeric hairpin RNA was constructed by PCR amplifying FucT and XylT forward and reverse gene fragments from *Lemna minor* (8627) cDNA and sequentially cloning them into pT7blue (EMD Biosciences) creating plasmid XF02 in T7-4. The FucT forward gene fragment was amplified with DNA primers BLX 686 (5'-ATGGTCGACTGCTGCTGGTGCTC TCAAC-3') (SEQ ID NO:22) and BLX690 (5'- ATGTCTAGAATG CAGCAGCAAGTGCACC-3') (SEQ ID NO:23) producing a 620 bp product with terminal SalI (5') and XbaI (3') cloning sites. The XylT forward gene fragment was amplified with DNA primers BLX 700 (5'-ATGACTAGTTGC GAAGCCTACTTCGGCAACAGC3') (SEQ ID NO:24) and BLX694 (5'-

ATGGGATCCGAATCTCAAGA ACAACTGTCG-3') (SEQ ID NO:25) producing a 1144 bp product with terminal SpeI (5') and BamHI (3') cloning sites. The XylT reverse gene fragment was amplified with DNA primers BLX 695 (5'-ATGGGTACCTGCGAAGCCTACTTCGGCAA CAGC-3') (SEQ ID NO:26) and BLX696 (5'- ATGGGA TCCACTGGCTGGGAGAAGTTCTT-3') (SEQ ID NO:27) producing a 644 bp product with terminal BamHI (5') and KpnI (3') cloning sites. The FucT reverse gene fragment was amplified with DNA primers BLX 691 (5'-ATGGAGCTCTGCTGCTGGTGCT CTCAAC-3') (SEQ ID NO:28) and BLX692 (5'-ATGGGTACCATGCAGCAGCAAGTGCACC-3') (SEQ ID NO:29) producing a 620 bp product with terminal KpnI (5') and SacI (3') cloning sites.

Independent expression cassettes containing promoter, gene of interest, and Nos terminator were created for the optimized MDX-060 H and L chains and the chimeric RNAi. Expression cassettes were cloned into a modification of the *Agrobacterium* binary vector pBMSP3 (obtained from Dr. Stan Gelvin, Purdue University) with the appropriate restriction sites. The H chain was fused to the modified chimeric octopine and mannopine synthase promoter with *Lemna gibba* 5' RbcS leader (Gasdaska *et al.* (2003) *Bioprocessing J.* 50-56). The L-chain was fused to the high expression, constitutive *Lemna minor* polyubiquitin promoter (LmUbq). The chimeric RNAi cassette, taken from plasmid XF02 in T7-4, was fused to the high expression, constitutive *Spirodela polyrhiza* polyubiquitin promoter (SpUbq). The three expression cassettes were cloned into the modified pBMSP3 binary vector in tandem orientation creating plasmid MDXA04.

#### *Transformation and plant line screening.*

Using *Agrobacterium tumefaciens* C58Z707 (Hepburn *et al.* (1985) *J. Gen. Microbiol.* 131:2961-2969), transgenic plants representing individual clonal lines were generated from rapidly growing *Lemna minor* nodules according to the procedure of Yamamoto *et al.* (Yamamoto *et al.* (2001) *In vitro Cell. Dev. Biol.* 37). For transgenic screening, individual clonal lines were preconditioned for 1 week at 150 to 200  $\mu\text{mol m}^{-2}\text{s}^{-2}$  in vented plant growth vessels containing SH media (Schenk and Hildenbrandt (1972) *Can. J. Botany* 50:199-204) without sucrose. Fifteen to twenty preconditioned fronds were then placed into vented containers containing fresh

SH media, and allowed to grow for two weeks. Tissue and media samples from each line were frozen and stored at -70°C.

*ELISA analysis of mAb produced in Lemna.*

5            *Lemna* tissue (100 mg) was homogenized using a FastPrep FP120 bead mill (Thermo Electron Corporation). Supernatants were diluted to 1 µg/mL and assayed using the IgG Quantitation ELISA kit (Bethyl Laboratories). For the assay, microtiter plates were coated with goat anti-human IgG at a concentration of 10 µg/mL, and mAb was detected by horseradish peroxidase (HRP)-conjugated goat anti-human IgG  
10            diluted 1:100,000. Standard curves were created with Human Reference IgG supplied with the ELISA kit. The sensitivity of the ELISA was 7.8 ng/mL. All samples were analyzed in duplicate.

15            *Preparation of Lemna microsomal membranes and assaying for core β-1,2-xylosyltransferase and α-1,3-fucosyltransferase activities.*

*Lemna* tissue (100 mg) from each line was homogenized in 1 mL of cold homogenization buffer (50 mM 4-[2-hydroxyethyl]-1-piperazineethanesulfonic acid [HEPES], pH 7.5, 0.25 M sucrose, 2 mM ethylenediaminetetraacetic acid [EDTA], 1 mM 1,4-dithiothreitol [DTT]) for 40 s in a FastPrep FP120 bead mill (Thermo  
20            Electron Corporation, Waltham, MA). The homogenate was centrifuged at 1,000 g for 5 min at 4°C. The supernatant was removed and centrifuged at 18,000g for 90 min at 4°C. The resulting pellet was resuspended in 20 µL of cold reaction buffer (0.1 M 2-[4-morpholino] ethanesulfonic acid [Mes], pH 7.0, 0.1% [v/v] Triton X-100, 10 mM MnCl<sub>2</sub>) and kept on ice or stored at -80°C until use.

25            Core β-1,2-xylosyltransferase and α-1,3-fucosyltransferase activities were measured simultaneously in 4 µL of microsomal membranes prepared from each RNAi line by incubating with 125 mM GlcNAc, 6.25 mM UDP-Xyl, 6.25 mM GDP-Fuc, 12.5 mM MnCl<sub>2</sub>, and 1.5 nmol of GnGn-dabsyl-peptide acceptor for 2 h at 37°C as described previously (Leiter *et al.* (1999) *J. Biol. Chem.* 274:21830-21839). The  
30            reaction was terminated by a brief centrifugation and incubation at 4°C and the products were analyzed by positive reflectron mode MALDI-TOF MS.



*Purification of MDX-060 LEX and LEX<sup>Opt</sup> mAbs.*

Plant tissue was homogenized with 50 mM sodium phosphate, 0.3 M sodium chloride, and 10 mM EDTA, pH 7.2 using a Silverson high shear mixer. The homogenate was acidified to pH 4.5 with 1 M citric acid, and centrifuged at 7,500g for 30 min at 4°C. The pH of the supernatant was adjusted to pH 7.2 with 2 M Tris, prior to filtration using 0.22 µm filters. The material was loaded directly on mAbSelect SuRe protein A resin (GE Healthcare) equilibrated with a solution containing 50 mM sodium phosphate, 0.3 M sodium chloride, and 10 mM EDTA, pH 7.2. After loading, the column was washed to baseline with the equilibration buffer followed by an intermediate wash with 5 column volumes of 0.1 M sodium acetate, pH 5.0. Bound antibody was eluted with 10 column volumes of 0.1 M sodium acetate, pH 3.0. The protein A eluate was immediately neutralized with 2 M 2-amino-2-[hydroxymethyl]-1,3-propanediol (Tris). For aggregate removal, the protein A eluate was adjusted to pH 5.5 and applied to a ceramic hydroxyapatite type I (Bio-Rad) column equilibrated with 25 mM sodium phosphate, pH 5.5. After washing the column with 5 column volumes of equilibration buffer, the antibody was eluted in a single step-elution using 0.25 M sodium phosphate, pH 5.5. Fractions containing antibody by A<sub>280</sub> were pooled and stored at -80°C.

Tissue extract and protein A flow through samples were prepared for SDS-PAGE under reducing and non-reducing conditions by addition of 2x SDS sample buffer ± 5% [v/v] 2-mercaptoethanol. Protein A eluate and hydroxyapatite eluate samples were diluted to a protein concentration of 0.5mg/mL followed by addition of 2x SDS sample buffer ± 5% [v/v] 2-mercaptoethanol. Samples were incubated at 95°C for 2 minutes prior to electrophoresis using 4-20% Tris-Glycine gradient gels (Invitrogen, Carlsbad, CA). Mark12 Molecular weight markers (Invitrogen) and a MDX-060 reference standard were included on the gels. Gels were stained with Colloidal Blue stain.

*Purification of N-linked glycans.*

Protein A purified monoclonal antibodies (1 mg) from wild-type and RNAi *Lemna* plant lines were dialyzed extensively against water and lyophilized to dryness. Samples were resuspended in 100 µL of 5% (v/v) formic acid, brought to 0.05 mg/ml pepsin, and incubated at 37°C overnight. The samples were heat inactivated at 95°C

for 15 min and dried. Pepsin digests were resuspended in 100  $\mu$ L of 100 mM sodium acetate, pH 5.0 and incubated with 1 mU of *N*-glycosidase A at 37°C overnight. The released *N*-glycans were isolated using 4 cc Carbograph SPE columns (Packer *et al.* (1998) *Glycoconj. J.* 19:737-747) and dried.

5           Dried *N*-glycans were further purified using 1 cc Waters Oasis MCX cartridges. Columns were prepared by washing with 3 column volumes of methanol followed by 3 column volumes of 5% (v/v) formic acid. *N*-glycans, resuspended in 1 mL of 5% (v/v) formic acid, were loaded onto the prepared columns. The unbound fraction as well as 2 additional column volume washes of 5% (v/v) formic acid were  
10           collected, pooled, and dried.

*Derivatization of oligosaccharides with 2-aminobenzoic acid (2-AA).*

Purified *N*-glycans or maltooligosaccharides were labeled with 2-AA and purified using 1 cc Waters Oasis HLB cartridges according to Anumula and Dhume,  
15           1998 (Anumula and Dhume (1998) *Glycobiology* 8:685-694). Labeled *N*-glycans and maltooligosaccharides were resuspended in 50  $\mu$ L of water and analyzed by negative mode MALDI-TOF MS and NP-HPLC-QTOF MS.

*MALDI-TOF Mass Spectrometry.*

20           MALDI-TOF MS was conducted using a Waters MALDI Micro MX (Millford, MA). Analysis of  $\beta$ -1,2-xylosyltransferase/ $\alpha$ -1,3-fucosyltransferase reaction products was conducted by mixing 0.5  $\mu$ L of each reaction supernatant with 0.5  $\mu$ L of 10 mg/mL CHCA in 0.05% (v/v) TFA, 50% (v/v) acetonitrile on a target plate. Xylosylated ( $[M+H]^+ = 2192.85$  Da) or fucosylated ( $[M+H]^+ = 2206.87$  Da)  
25           GnGn-dabsyl-peptide products were detected in positive reflectron mode. Ion counts of 200 combined spectra from each sample were normalized against that of  $\beta$ -1,4-galactosylated, GnGn-dabsyl-peptide ( $[M+H]^+ = 2222.87$  Da) present as a contaminant (<5%) in the original GnGn-dabsyl-peptide mixture from EMD Biosciences.

30           2-AA labeled *N*-glycans or maltooligosaccharides (0.5  $\mu$ L) were diluted with water, mixed with 0.5  $\mu$ L of 10 mg/ml DHB matrix in 70% (v/v) acetonitrile, spotted onto a target plate and analyzed in negative reflectron mode.

*NP-HPLC-Q-TOF MS analysis of 2-AA labeled N-glycans.*

2-AA labeled *N*-glycans or maltooligosaccharides were brought to 80% (v/v) acetonitrile and separated on a Waters 2695 HPLC system fitted with a TSK-Gel Amide-80 (2 mm x 25 cm, 5  $\mu$ m) column (Tosoh Biosciences, Montgomeryville, PA).  
5 2-AA labeled carbohydrates were detected and analyzed using a Waters 2475 fluorescence detector (230 nm excitation, 425 nm emission) and a Waters Q-TOF API US quadropole-time of flight (QTOF) mass spectrometer fitted on-line with the HPLC system.

Separations were conducted at 0.2 mL/min, 40°C, using 10 mM ammonium acetate, pH 7.3 (solvent A) and 10 mM ammonium acetate, pH 7.3, 80% (v/v) acetonitrile (solvent B). Sample elution was carried out at 0% A isocratic for 5 min, followed by a linear increase to 10% A at 8 min, and a linear increase to 30% A at 48 min. The column was washed with 100% A for 15 min and equilibrated at 0% A for 15 min prior to the next injection.

15 QTOF analysis was conducted in negative ion mode with source and desolvation temperatures of 100°C and 300°C, respectively, and capillary and cone voltages of 2,100 and 30 V, respectively. Mass spectra shown are the result of combining  $\geq 40$  individual scans per labeled *N*-glycan.

20 *Monosaccharide analysis by HPAEC-PAD.*

mAb samples (200  $\mu$ g) were subjected to acid hydrolysis using 2 N TFA at 100°C for 3 h. Samples were dried by vacuum centrifugation at ambient temperature and reconstituted in 150  $\mu$ L water prior to analysis by HPAE-PAD (Dionex). An aliquot (25  $\mu$ L) of the reconstituted sample was applied to a CarboPac PA10 column  
25 (4 x 250 mm) with a pre-column Amino Trap (Dionex). Separation of monosaccharides was accomplished with a mobile phase of 3.5 mM KOH, using an EG40 eluent generator. Monosaccharide peak identity and relative abundance were determined using monosaccharide standards.

30 *Thermal stability of mAb.*

A MicroCal (Northampton, MA) VP-Capillary differential scanning calorimetry (DSC) instrument was used to determine thermal stability of glycol-optimized and wild-type mAbs. Purified mAb samples were dialyzed in 20 mM

NaH<sub>2</sub>PO<sub>4</sub>, pH 7.4, 150 mM NaCl (PBS) overnight. Thermal denaturation data was collected by heating the samples at a concentration of 300 µg/mL from 35 to 95 °C at a scan rate of 1°C/min using PBS as the reference buffer. The feedback and gain were set to low. The baseline-corrected and normalized data was fit to a non-2-state model using Origin v7.0 software.

*FcR binding activity of mAb.*

The experiment was conducted using a BIACORE (Biacore AB, Uppsala, Sweden) instrument using surface plasmon resonance technology. mAbs, 2 µg/mL, were captured on the antigen coated surface (recombinant human CD30). Several concentrations of both the Val<sup>158</sup> and Phe<sup>158</sup> allotypes of FcRγIIIa, starting from 6 µM, were flowed over the captured antibodies for 3 min. The binding signal as a function of FcRγIIIa was fit to a one-site binding model using GraphPad Prism (v4.01) software to obtain the K<sub>D</sub> values. HBS-EP buffer (10 mM HEPES, 0.15 M NaCl, 3 mM EDTA and 0.005% (v/v) P20 at pH 7.4) was used throughout the experiment. Binding of the mAbs to buffer or FcRγIIIa to blank surfaces were used as negative controls.

*Assay for antigen binding affinity.*

CD30-expressing L540 cells (DSMZ Cell Culture Collection # ACC 72) were used as antigen positive cells to assay for binding. Aliquots of 2 x 10<sup>5</sup> cells/well were incubated for 30 min at 4°C with 100 µL of primary antibody at the indicated concentrations. Cells were washed twice in PBS with 2% (v/v) fetal bovine serum (FBS) before addition of goat anti-human mAb, FITC-labeled secondary antibody (Jackson ImmunoResearch, West Grove, PA) at 1:500 dilution in 100 µL/well for 30 min at 4°C. Cells were washed twice in PBS with 2% (v/v) FBS and assayed by flow cytometry using a FACS Calibur instrument (Becton Dickinson, Franklin Lakes, NJ). EC<sub>50</sub> values of MDX-060 CHO, LEX and LEX<sup>Opt</sup> mAb binding to CD30 on L540 cells were determined from binding curves utilizing GraphPad Prism 3.0 software.

*ADCC assay.*

Human peripheral blood mononuclear effector cells were purified from heparinized whole blood by standard Ficoll-Paque separation. Cells (2 x 10<sup>6</sup>) were

washed in PBS and sent for genotyping. The remaining effector cells were then resuspended at  $1 \times 10^6$  cells/mL in RPMI 1640 medium containing 10% (v/v) FBS and 50 U/mL of human IL-2 (Research Diagnostics, Concord, MA) and incubated overnight at 37°C. The effector cells were washed once in culture medium and resuspended at  $1 \times 10^7$  cells/mL prior to use. L540 target cells at  $1 \times 10^6$  cells/mL in RPMI 1640 medium containing 10% (v/v) FBS and 5 mM probenecid were labeled with 20  $\mu$ M BATDA (bis(acetoxymethyl)2,2':6',2"-terpyridine-6,6"-dicarboxylate) for 20 min at 37°C. Target cells were washed three times in PBS supplemented with 20 mM HEPES and 5 mM probenecid, resuspended at  $1 \times 10^5$  cells/mL and added to effector cells in 96-well plates ( $1 \times 10^4$  target cells and  $5 \times 10^5$  effector cells/well) at a final target to effector ratio of 1:50. Maximal release was obtained by incubation of target cells in 3% (v/v) Lysol and spontaneous release obtained by incubation in cell culture medium alone. After 1 h incubation at 37°C, 20  $\mu$ L of supernatant was harvested from each well and added to wells containing 180  $\mu$ L of Europium solution. The reaction was read with a Perkin Elmer Fusion Alpha TRF reader using a 400  $\mu$ sec delay and 330/80, 620/10 excitation and emission filters respectively. The counts per second were plotted as a function of antibody concentration and the data was analyzed by non-linear regression, sigmoidal dose response (variable slope) using GraphPad Prism 3.0 software. The percent specific lysis was calculated as: (experimental release – spontaneous release)/(maximal release – spontaneous release) x 100. In all studies, human mAb1 isotype control was included and compared to MDX-060 CHO, LEX, and LEX<sup>Opt</sup> mAbs. Other controls included target and effector cells with no mAb, target cells with no effector cells and target and effector cells in the presence of 3% (v/v) Lysol.

25

## RESULTS

### *Expression of MDX-060 mAb in the LEX system.*

MDX-060 is an anti-CD30 antibody (formally known as 5F11) being developed for the treatment of Hodgkins lymphoma and anaplastic large cell lymphoma (Borchmann *et al.* (2003) *Blood* 102:3737-3742). Two binary vectors were constructed for the expression of MDX-060 in the LEX system. Expression vector MDXA01 contained codon optimized genes encoding heavy (H) and light (L) chains of MDX-060 while vector MDXA04 contained genes encoding H and L chains

30



in addition to a chimeric FucT/XylT RNAi gene (Figure 39). Independent transgenic lines were generated for both the MDXA01 (165 lines) and MDXA04 (195 lines) expression vectors. For simplicity, MDXA01 derived mAbs (wild-type *N*-glycosylation), and MDXA04 derived mAbs (containing the FucT/XylT RNAi construct) will be referred to as MDX-060 LEX and MDX-060 LEX<sup>Opt</sup>, respectively, in the discussions below.

Transgenic plant lines were first screened for mAb expression with an IgG ELISA. LEX<sup>Opt</sup> lines with high levels of mAb expression were assayed further for FucT and XylT activity. Transferase activities in the majority of the high expressing MDX-060 LEX<sup>Opt</sup> lines were reduced to levels of the negative control indicating effective silencing in the majority of the assayed lines (Figure 40). MDX-060 LEX<sup>Opt</sup> lines did not exhibit any morphological or growth differences compared to wild-type *Lemna* plants (data not shown).

Thermal stabilities of the MDX-060 CHO, LEX, and LEX<sup>Opt</sup> mAbs were determined using differential scanning calorimetry (DSC). All three mAbs exhibited similar melting curve kinetics (data not shown) and melting transition point temperatures (Table 2 below), further demonstrating the structural integrity of the *Lemna*-produced MDX-060 LEX and LEX<sup>Opt</sup> mAbs compared to the MDX-060 CHO mAb. SDS-PAGE analysis under non-reducing (Figure 41A) and reducing conditions (Figure 41B) showed that mAbs from the MDX-060 LEX<sup>Opt</sup> and MDX-060 CHO lines had similar protein profiles with the mAb appearing as the major component in the protein extract.

Table 2. Comparison of the thermal stabilities of MDX-060 CHO, MDX-060 LEX, and glyco-optimized MDX-060 LEX<sup>Opt</sup> mAbs by differential scanning calorimetry (DSC).

Antibody	T <sub>m1</sub> (°C)	T <sub>m2</sub> (°C)	T <sub>m3</sub> (°C)
MDX-060 CHO	72	75	84
MDX-060 LEX	71	75	84
MDX-060 LEX <sup>Opt</sup>	72	76	84

*N-glycan structures of MDX-060 CHO, LEX, and LEX<sup>Opt</sup> mAbs.*

The *N*-glycan profiles of recombinant MDX-060 CHO, MDX-060 LEX, and MDX-060 LEX<sup>Opt</sup> derived mAbs were determined by negative reflectron mode MALDI-TOF MS and normal phase (NP) HPLC-QTOF MS. The structures of *N*-glycans referred to in the following discussion are shown in Figure 53.

MALDI-TOF MS analysis of *N*-glycans from MDX-060 CHO lines indicated the presence of four major *N*-glycans with *m/z* values corresponding to 2-AA labeled GnGnF<sup>6</sup> (nomenclature derived from <http://www.proglycan.com>), Man5, GnA<sub>iso</sub>F<sup>6</sup>, and AAF<sup>6</sup> (Figure 42). NP-HPLC separated the GnA<sub>iso</sub>F<sup>6</sup> *N*-glycan into its two isoforms (Gal attached to the  $\alpha$ -1,6-Man or  $\alpha$ -1,3-Man arm) bringing the total number of major *N*-glycans found on MDX-060 CHO to five (Figure 43). MS/MS fragmentation of the peaks was not conducted to confirm the identity of each isoform; however, the higher abundance of the earlier peak suggested that Gal was attached to the  $\alpha$ -1,6-Man arm of this *N*-glycan (Shinkawa *et al.* (2003) *J. Biol. Chem.* 278:3466-3473; Zhu *et al.* (2005) *Nat. Biotechnol.* 23:1159-1169). On-line negative mode QTOF MS analysis gave *m/z* values corresponding to doubly charged GnGnF<sup>6</sup>, Man5, GnA<sub>iso</sub>F<sup>6</sup> (both isoforms), and AAF<sup>6</sup>, confirming the MALDI-TOF MS results (Table 3 below). Peak integration of the fluorescent trace revealed that GnGnF<sup>6</sup>, Man5, AGnF<sup>6</sup>, GnAF<sup>6</sup>, and AAF<sup>6</sup> constituted 50.8, 2.5, 26.1, 10.7 and 6.8%, respectively, of the total *N*-glycan structures from MDX-060 CHO. The remaining 3.1% of *N*-glycans were found to be a mixture of GnGn, GnM<sub>iso</sub>F<sup>6</sup>, GnM<sub>iso</sub>, and MM with no structure higher than 1.2% of the total (data not shown).

Table 3. Summary of observed MALDI-TOF and QTOF MS masses of the major 2-AA labeled *N*-glycans from MDXA-060 mAbs produced by CHO cells (CHO), wild-type *Lemna* (LEX) or glyco-optimized *Lemna* lines expressing the RNAi construct (LEX<sup>Opt</sup>).

<i>N</i> -glycan name <sup>a</sup>	Proposed Structure <sup>b</sup>	Theoretical m/z		Observed MALDI-TOF <sup>c</sup>	Observed Q-TOF <sup>c</sup>	% Peak Area <sup>c</sup>
		[M-H] <sup>-</sup>	[M-2H] <sup>2-</sup>	[M-H] <sup>-</sup>	[M-2H] <sup>2-</sup>	
<b>CHO</b>						
GnGnF <sup>6</sup> -2AA		1582.590	790.7911	1582.455	790.7825	50.8
Man5-2AA		1354.479	676.7436	1354.392	676.7343	2.50
GnA <sub>iso</sub> F <sup>6</sup> -2AA		1744.642	871.8175	1744.492	871.7970	36.8
AAF <sup>6</sup> -2AA		1906.695	952.8438	1906.567	952.8181	6.80
<b>LEX</b>						
GnGn-2AA		1436.532	717.7622	1436.549	717.7894	8.40
GnGnX-2AA		1568.574	783.7833	1568.581	783.8150	17.2
GnGnXF <sup>3</sup> -2AA		1714.632	856.8122	1714.615	856.853	67.4
<b>LEX<sup>Opt</sup></b>						
GnGn-2AA		1436.532	717.7622	1436.523	717.7993	95.8

<sup>a</sup>*N*-glycan names are based on Proglycan (<http://www.proglycan.com>) nomenclature. 2AA, 2-aminobenzoic acid.

<sup>b</sup>The symbols of the proposed *N*-glycan structures are as follows: , *N*-acetylglucosamine; , mannose; , galactose; , xylose; ,  $\alpha$ -1,3-fucose; ,  $\alpha$ -1,6-fucose; , 2-aminobenzoic acid.

<sup>c</sup>The m/z values and the % peak area of each *N*-glycan structure were obtained from Figure 2.

5

MALDI-TOF MS analysis of wild-type MDX-060 LEX mAb revealed the presence of three major species with m/z values corresponding to GnGnXF<sup>3</sup>, GnGnX

and GnGn (Figure 42). NP-HPLC-QTOF MS analysis showed three major fluorescent peaks with m/z values corresponding to doubly charged GnGnXF<sup>3</sup>, GnGnX and GnGn, again confirming the MALDI-TOF MS results (Figure 43; Table 3). Integration of the fluorescent peaks indicated that GnGnXF<sup>3</sup>, GnGnX and GnGn constituted 67.4, 17.2 and 8.4%, respectively, of the total *N*-glycans derived from the MDX-060 LEX mAb. The remaining 7% of *N*-glycans were determined to be a mixture of MM, GnM<sub>iso</sub>, MMXF<sup>3</sup>, GnGnF<sup>3</sup>, GnM<sub>iso</sub>XF<sup>3</sup>, Man6, Man7, Gn(FA)<sub>iso</sub>XF<sup>3</sup>, Man8 and Man9 with no *N*-glycan greater than 2% of the total. Similar results were seen with mAbs isolated from two independently transformed MDX-060 LEX lines (data not shown). The simple array of *N*-glycans on LEX mAbs demonstrated here provides an amenable starting point for glyco-optimization.

In contrast to the MDX-060 LEX mAb, *N*-glycans from the MDX-060 LEX<sup>Opt</sup> mAb possessed GnGn as the major *N*-glycan species by both MALDI-TOF MS and NP-HPLC-QTOF MS analysis (Figures 42 and 43; Table 3). GnGn comprised 95.8% of the total *N*-glycans with the remaining 4.2% of *N*-glycans determined to be MM, GnM<sub>iso</sub>, GnA<sub>iso</sub>, Man6, Man7 and Man8 with no one structure greater than 1.2% of the total *N*-glycans. None of the LEX<sup>Opt</sup> *N*-glycans contained fucose (Fuc) or xylose (Xyl). These results demonstrated that co-expression of an RNAi construct targeting *Lemna* FucT and XylT resulted in the complete elimination of Fuc and Xyl-containing *N*-glycans from MDX-060 LEX<sup>Opt</sup> mAbs and produced highly homogeneous mAb glycoforms. The same results were obtained for MDX-060 LEX<sup>Opt</sup> mAb harvested from an independent transgenic line (line 225) comprising the MDXA04 expression vector, at a different growth scale (300 g tissue for transgenic line 225 versus 1 g tissue for transgenic line 52, which produced the MDX-060 LEX<sup>Opt</sup> mAb having the *N*-glycan profile shown in Figures 42 and 43. Unlike mammalian cell culture systems where *N*-glycan heterogeneity can change with culture conditions, growth scale, and growth period (Kanda *et al.* (2006) *Biotechnol. Bioeng.* 94:680-688), the glycan uniformity observed with LEX<sup>Opt</sup> mAbs was shown to be consistent under a variety of growth conditions and scales.

The absence of Fuc or Xyl on MDX-060 LEX<sup>Opt</sup> mAb *N*-glycans was further confirmed by monosaccharide analysis (Table 4 below). Monosaccharides were released from MDX-060 CHO, LEX and LEX<sup>Opt</sup> mAbs by acid hydrolysis and analyzed by high performance anion exchange chromatography (HPAEC) coupled to

pulsed amperometric detection (PAD). The monosaccharide ratios for Man and GlcNAc residues were similar for CHO and wild-type LEX mAbs and correlated well with expected values. LEX mAbs were significantly decreased in Gal and Fuc content and had a significant increase in Xyl when compared to CHO-derived mAbs.

5 Monosaccharide analysis of *Lemna* derived mAbs revealed that while Fuc and Xyl were present on wild-type LEX *N*-glycans, they were not detected on LEX<sup>Opt</sup> mAbs. Collectively, these results demonstrate that co-expression of an RNAi construct targeting *Lemna* XylT and FucT results in the elimination of Fuc and Xyl-containing *N*-glycans from MDX-060 LEX<sup>Opt</sup> mAbs and produce highly homogeneous mAb

10 glycoforms. The robustness of this glyco-optimization strategy has been confirmed with multiple independent plant lines expressing the MDX-060 LEX<sup>Opt</sup> mAb as well as with other mAbs expressed in the LEX System, for example, the mAbI monoclonal antibody discussed in the examples herein above. In these subsequent transformations, glyco-optimized mAb expression levels up to 6% of total soluble

15 protein (TSP) have been obtained.

Table 4. Monosaccharides released from MDX-060 CHO, LEX and LEX<sup>Opt</sup> mAbs by acid hydrolysis and analyzed by HPAEC-PAD. The monosaccharide content from each mAb was determined by normalizing against carbohydrate controls.

20

Monosaccharide	MDX-060 CHO pmol (% total)	MDX-060 LEX pmol (% total)	MDX-060 LEX <sup>Opt</sup> pmol (% total)
Fuc	254 (20)	232 (13)	0
GlcNAc	605 (47)	773 (45)	1,003 (67)
Gal	75 (6)	0	0
Man	355 (27)	491 (29)	501 (33)
Xyl	0	226 (13)	0
Total	1,289 (100)	1,722 (100)	1,504 (100)



*Functional activity of MDX-060 CHO, LEX and LEX<sup>Opt</sup> mAbs .*

Antigen binding properties of the MDX-060 CHO, MDX-060 LEX, and MDX-060 LEX<sup>Opt</sup> mAbs were determined using CD30 expressing L540 cells. All three mAbs had nearly identical binding curves (Figure 43). EC<sub>50</sub> concentrations were determined to be 0.180 µg/mL, 0.227 µg/mL, and 0.196 µg/mL for MDX-060 CHO, LEX, and LEX<sup>Opt</sup>, respectively (Figure 44), indicating that antigen binding for all three mAbs were similar.

Fc-receptor-mediated effector cell function has been shown to be important for the *in vivo* activity of many therapeutic mAbs. Since the FcR expressed on NK cells and macrophages responsible for ADCC activity is FcγRIIIa, binding of the various mAbs to this receptor was compared. FcR binding of CHO, LEX and LEX<sup>Opt</sup> mAbs was determined by equilibrium binding of the mAbs with effector cells expressing two different human FcγRIIIa allotypes (Phe<sup>158</sup> or Val<sup>158</sup>). MDX-060 LEX had a 1.7-fold increase in FcγRIIIaPhe<sup>158</sup> and a 0.4-fold decrease in FcγRIIIaVal<sup>158</sup> binding compared to the CHO-derived mAb, demonstrating that receptor binding for CHO and LEX mAbs were similar. In contrast, LEX<sup>Opt</sup> mAbs had a 27 and 15-fold higher affinity for FcγRIIIaPhe<sup>158</sup> and FcγRIIIaVal<sup>158</sup>, respectively, than CHO mAbs (Figure 45). These results suggested that RNAi silencing of the *Lemna* FucT and XylT activities in LEX<sup>Opt</sup> lines produced mAbs with enhanced FcR binding.

ADCC activities of the CHO, LEX and LEX<sup>Opt</sup> mAbs were determined by incubating mAbs with either homozygous (FcγRIIIaPhe<sup>158</sup>) or heterozygous (FcγRIIIaPhe/Val<sup>158</sup>) human effector cells and BATDA (bis(acetoxymethyl)2,2':6',2''-terpyridine-6,6''-dicarboxylate) labeled L540 target cells (Figure 45). MDX-060 LEX mAbs (31%) had the same maximal percent cell lysis as CHO mAbs (31%) using heterozygous FcγRIIIaPhe/Val<sup>158</sup> human effector cells (Figure 46) with similar EC<sub>50</sub> values. Maximal percent cell lysis for LEX mAbs (27%) was slightly increased compared to CHO mAbs (15%) using homozygous FcγRIIIa Phe/Phe<sup>158</sup> effector cells. Importantly, LEX<sup>Opt</sup> mAbs had significantly increased ADCC activity compared to MDX-060 CHO and LEX mAbs, irrespective of the donor genotype. This was assessed by both an increase in potency and efficacy. Maximal percent lysis for MDX-060 Lex<sup>Opt</sup> was 45% for both experiments, while the EC<sub>50</sub> value was 3 to 5 times lower than MDX-060 LEX and MDX-060 CHO mAbs, respectively, for FcγRIIIa Val/Phe<sup>158</sup> effector cells and 2 to 3 times lower

for the FcγRIIIa Phe/Phe<sup>158</sup> effector cells. These results demonstrate that removal of Fuc and Xyl-containing *N*-glycans from MDX-060 LEX<sup>Opt</sup> mAbs caused an enhancement in ADCC activity and hence can improve their therapeutic potential.

5 *RP-HPLC-Q-TOF MS analysis of intact IgG for MDX-060 LEX and MDX-060 LEX<sup>Opt</sup>.*

Protein A purified IgG's (50 μg) were desalted using the Waters 2695 HPLC system fitted with a Poros R1-10 column (2 mm x 30 mm; Applied Biosystems).  
 10 IgG's were detected and analyzed using a Waters 2487 dual wavelength UV detector (280 nm) and the Waters Q-TOF API US. Separations were conducted at 0.15 mL/min, 60°C, using 0.05% (v/v) trifluoroacetic acid (TFA; solvent A) and 0.05% (v/v)TFA, 80% (v/v) acetonitrile (solvent B). Sample elution was carried out using a linear increase from 30 to 50% B for 5 min, an increase to 80% B for 5 min. The  
 15 solvent ratio remained at 80% B for an additional 4 min, followed by a wash with 100% B for 1 min and equilibration of the column with 30% B for 15 min prior to the next run.

Q-TOF analysis was conducted in positive ion mode with source and desolvation temperatures of 100°C and 300°C, respectively, and capillary and cone  
 20 voltages of 3.0 and 60 V, respectively. Data are the result of combining ≥100 individual scans and deconvolution to the parent mass spectrum using MaxEnt 1.

See also Triguero *et al.* (2005) *Plant Biotechnol. J.* 3: 449-457; Takahashi *et al.* (1998) *Anal. Biochem.* 255: 183-187; Dillon *et al.* (2004) *J. Chromatogr. A.* 1053: 299-305.

25 Figure 47 shows intact mass analysis of the MDX-060 LEX mAb compositions produced in wild-type *L. minor* comprising the MDXA01 construct. When XylT and FucT expression are not suppressed in *L. minor*, the recombinantly produced MDX-060 LEX mAb composition comprises at least 7 different glycoforms, with the G0XF<sup>3</sup> glycoform being the predominate species present. Note  
 30 the absence of a peak representing the G0 glycoform.

Figure 48 shows glycan mass analysis of the heavy chain of the MDX-060 LEX mAb produced in wild-type *L. minor* comprising the MDXA01 construct. When XylT and FucT expression are not suppressed in *L. minor*, the predominate *N*-glycan

species present is G0XF<sup>3</sup>, with additional major peaks reflecting the G0X species. Note the minor presence of the G0 glycan species.

Figure 49 shows intact mass analysis of the MDX-060 LEX<sup>Opt</sup> mAb compositions produced in transgenic *L. minor* comprising the MDXA04 construct. When XylT and FucT expression are suppressed in *L. minor*, the intact mAb composition contains only G0 *N*-glycans. In addition, the composition is substantially homogeneous for the G0 glycoform (peak 2), wherein both glycosylation sites are occupied by the G0 *N*-glycan species, with two minor peaks reflecting trace amounts of precursor glycoforms (peak 1, showing mAb having an Fc region wherein the C<sub>H</sub>2 domain of one heavy chain has a G0 glycan species attached to Asn 297, and the C<sub>H</sub>2 domain of the other heavy chain is unglycosylated; and peak 3, showing mAb having an Fc region wherein the Asn 297 glycosylation site on each of the C<sub>H</sub>2 domains has a G0 glycan species attached, with a third G0 glycan species attached to an additional glycosylation site within the mAb structure).

Figure 50 shows glycan mass analysis of the heavy chain of the MDX-060 LEX<sup>Opt</sup> mAb produced in transgenic *L. minor* comprising the MDXA04 construct. When XylT and FucT expression are suppressed in *L. minor*, the only readily detectable *N*-glycan species attached to the Asn 297 glycosylation sites of the C<sub>H</sub>2 domains of the heavy chains is G0.

## DISCUSSION

Glyco-optimization of MDX-060 was accomplished by co-expression with an RNAi cassette aimed at silencing the endogenous *Lemna* FucT and XylT genes. This simultaneous silencing of both FucT and XylT genes was achieved using a single RNAi transcript. The absence of Fuc and Xyl on the LEX<sup>Opt</sup> mAb was confirmed by MALDI-TOF, NP-HPLC-QTOF MS, and monosaccharide analysis of *N*-glycans purified from the MDX-060 LEX<sup>Opt</sup> mAb. These analyses corroborate the lack of transferase activity observed in microsomal membranes. Importantly, >95% of the *N*-glycans released from LEX<sup>Opt</sup> mAbs were of a single structure, GnGn, indicating that this strategy had the added benefit of producing mAbs with a homogeneous *N*-glycan profile. MDX-060 LEX and LEX<sup>Opt</sup> mAbs were found to be indistinguishable with regard to thermal stability and antigen binding compared to MDX-060 CHO.

Electrophoretic analysis was also found to be similar for all three mAbs. In fact, the only structural differences detected were in the mAb *N*-glycan profiles.

Without being bound by theory, the ability of the MDX-060 LEX<sup>Opt</sup> lines to produce mAbs with a single major *N*-glycan species may be based on the more uniform mAb glycoform distribution found in wild-type *Lemna*. *N*-glycans released from mAbs purified from wild-type tobacco (Fujiyama *et al.* (2006) *J. Biosci. Bioeng.* 101:212-218; Elbers *et al.* (2001) *Plant Physiol.* 126:1314-1322; Bakker *et al.* (2001) *Proc. Natl. Acad. Sci.* 98:2899-2904), alfalfa (Bardor *et al.* (2003) *Plant Biotechnol. J.* 1:451-462), and moss (Koprivova *et al.* (2003) *Plant Bio.* 5:582-591) show that mAb glycoform heterogeneity in plants with wild-type *N*-glycosylation can range from five (alfalfa) (Bardor *et al.* (2003) *Plant Biotechnol. J.* 1:451-462) to eight (tobacco) (Fujiyama *et al.* (2006) *J. Biosci. Bioeng.* 101:212-218) different major structures. MDX-060 LEX possesses only three major *N*-glycan structures (GnGn, GnGnX and GnGnXF<sup>3</sup>). This simple array of *N*-glycans on mAbs produced by wild-type *Lemna* may provide a more amenable starting point for glyco-optimization leading to greater homogeneity than that observed in other systems.

Fc-receptor mediated effector cell function has been shown to be important for the *in vivo* activity of many therapeutic mAbs. In this study, the ADCC activity of MDX-060 CHO, MDX-060 LEX, and MDX-060 LEX<sup>Opt</sup> mAbs was compared. Since the FcR expressed on NK cells and macrophages responsible for ADCC activity is FcγRIIIa, the binding of the various mAbs to this receptor was also compared. The results discussed above show that MDX-060 LEX<sup>Opt</sup> mAb has an increased binding affinity (15-25 fold) and maximal binding (4-5 fold) to FcγRIIIa as well as enhanced ADCC activity compared to MDX-060 CHO and MDX-060 LEX mAbs. The removal of α-1,6-linked Fuc from various mAbs produced in other expression systems has been shown previously to increase FcR binding and enhance ADCC function (Shinkawa *et al.* (2003) *J. Biol. Chem.* 278:3466-3473; Shields *et al.* (2002) *J. Biol. Chem.* 277:26733-26740; Niwa *et al.* (2004) *Clin. Cancer Res.* 10:6248-6255). The results presented herein suggest that removal of the α-1,3-linked Fuc from the MDX-060 LEX<sup>Opt</sup> mAbs has the same effect on mAb function as the removal of α-1,6-linked Fuc.

In this study, two naturally occurring polymorphic isoforms of FcγRIIIa at residue 158<sup>41</sup>, Val<sup>158</sup> and Phe<sup>158</sup>, were evaluated. MDX-060 LEX<sup>Opt</sup> shows a higher

binding affinity to FcγRIIIa-Val<sup>158</sup> compared to FcγRIIIa-Phe<sup>158</sup> as has been observed with other IgG1 mAbs (Shields *et al.* (2002) *J. Biol. Chem.* 277:26733-26740). The fact that an increase in binding with MDX-060 LEX<sup>Opt</sup> was observed with both isoforms is important since differential binding to Val<sup>158</sup> over Phe<sup>158</sup> was found to be predictive of the clinical and immunological responsiveness of certain patient groups receiving anti-CD20 treatment (Cartron *et al.* (2002) *Blood* 99:754-758; Weng *et al.* (2003) *J. Clin. Oncol.* 21:3940-3947; Weng *et al.* (2004) *J. Clin. Oncol.* 22:4717-4724). This increase in binding has been hypothesized to result in a higher percentage of patients responding to treatment that requires Fc functionality.

A similar increase in ADCC activity was also observed. In this study, the MDX-060 LEX<sup>Opt</sup> mAb showed an increase in cell lysis and a decrease in the EC<sub>50</sub> value, resulting in an increase in efficacy and potency when compared to MDX-060 CHO. This corresponds to a 20-fold increase in activity, determined by taking the maximum percent lysis of MDX-060 CHO and calculating the concentration of MDX-060 LEX<sup>Opt</sup> mAb giving rise to the same percent cell lysis. As with the FcγRIIIa binding study, the increase in ADCC activity was observed with both a homozygous FcγRIIIaPhe/Phe<sup>158</sup> and a heterozygous FcγRIIIa Phe/Val<sup>158</sup> effector cell donor. The results presented here suggest that removal of the α-1,3-linked Fuc from the MDX-060 LEX<sup>Opt</sup> mAbs has the same effect on mAb function as the removal of α-1,6-linked Fuc.

The robustness of this glyco-optimization strategy has been demonstrated with multiple independent *Lemna* plant lines expressing the MDX-060 LEX<sup>Opt</sup> mAb as well as with other mAbs expressed in the *Lemna* expression system (see, for example, Examples 2-4 above). Furthermore, there is no apparent difference in plant phenotype or growth rate compared with wild-type *Lemna* plants. Unlike mammalian cell culture systems where *N*-glycan heterogeneity can change with culture conditions, growth scale and growth period<sup>8</sup>, the glycan uniformity observed with LEX<sup>Opt</sup> mAbs has been shown to be consistent under a variety of growth conditions and scales (data not shown).

In conclusion, an RNAi strategy was used to produce a glyco-optimized anti-CD30 antibody in the *Lemna* expression system. The resulting mAb consists of a single, major *N*-glycan structure, without any evidence of the plant-specific Fuc and Xyl residues. In addition, the resulting optimized mAb has increased ADCC activity



and FcγRIIIa binding activity compared to a CHO-derived mAb. The homogeneous glycosylation profile obtained on mAbs produced in a *Lemna* expression system having this FucT+XylT gene knockout strategy makes it is possible to express these mAbs with increased production consistency.

5

Example 7: Scale-up Production of Glycan-Optimized mAbI in *Lemna minor*

*L. minor* transgenic line 24 comprising the mAbI04 construct of Figure 12 (providing for suppression of FucT and XylT) was generated in a manner similar to that described above. Following its generation, transgenic line 24 was continuously maintained by clonal culture, wherein periodically a subsample of the plant culture was transferred to fresh culture medium for further culturing. This transgenic line was analyzed for the *N*-glycosylation pattern of the recombinantly produced mAbI antibody following production scale-up from 1 g tissue up to 300 g (0.3 kg) tissue, and further production scale-up to 6.5 kg tissue. The process of scaling production up to 6.5 kg tissue occurred approximately 8 months after transgenic line 24 was generated. Results are shown in Figures 51A (MALDI-TOF analysis of *N*-glycans) and 51B (HPLC fluorescence analysis of *N*-glycans).

The glycosylation profile for the mAbI antibody produced by transgenic line 24 comprising the mAbI04 construct remained homogeneous with scale-up in production from 1 g tissue to 0.3 kg tissue, and further scale-up in production to 6.5 kg tissue, and thus was characterized by the presence of a single predominant peak corresponding to the GnGn (i.e., G0) glycan species. Thus, the homogeneity of the glycosylation profile in transgenic *L. minor* comprising the mAbI04 construct was consistently maintained with an approximately 6,500-fold increase in production scale (i.e., from 1 g up to 6.5 kg). Furthermore, the homogeneity of the glycosylation profile was consistently maintained in this transgenic line at 8 months following its generation.

These data demonstrate that the homogeneity of the glycosylation profile in transgenic *L. minor* comprising the mAbI04 construct remains consistent with at least a 6,500-fold increase in production scale. Furthermore, the homogeneity of the glycosylation profile in transgenic *L. minor* comprising the mAbI04 construct is maintained for at least 8 months after the transgenic line is generated. The homogeneity of the glycosylation profile would be expected to be maintained with



further increase in production scale, and thus, for example, would be expected to be maintained if production scale was increased by another 4-fold beyond 6.5 kg (e.g., scale-up from 6.5 kg to 26 kg). The homogeneity of the glycosylation profile would also be expected to be maintained with continuous clonal culture of the transgenic line well beyond 8 months after generation of the transgenic line.

Example 8: Glycosylation Pattern for Endogenous Glycoproteins in *Lemna minor* Lines Transgenic for mAbI04 RNAi Construct

The L40 protease is a representative endogenous glycoprotein produced in *L. minor*. In order to assess the impact of the mAbI04 RNAi construct (Figure 12) on glycosylation of endogenous proteins, the L40 protein was isolated from a *L. minor* line transgenic for the mAbI04 RNAi construct using benzamidine affinity chromatography. The *N*-glycosylation pattern for the isolated L40 protein was analyzed using MALDI-TOF analysis in the manner described above. Results are shown in Figure 52.

As can be seen from this analysis, suppression of FucT and XylT expression using the chimeric RNAi mAbI04 construct results in endogenous glycoproteins having a homogeneous glycosylation pattern consistent with that observed for recombinant glycoproteins. Thus, the heterogeneous *N*-glycan profile for the L40 glycoprotein isolated from *L. minor* having the wild-type glycosylation machinery is represented by a mixture of *N*-glycans species having the  $\beta$ 1,2-linked xylose residue, or both the  $\beta$ 1,2-linked xylose and core  $\alpha$ 1,3-linked fucose residues attached. In contrast, the homogeneous *N*-glycan profile for L40 isolated from *L. minor* transgenic for the mAbI04 RNAi construct is represented by a single predominant peak corresponding to the G0 glycan species, and is characterized by the absence of *N*-glycan species having the  $\beta$ 1,2-linked xylose or both the  $\beta$ 1,2-linked xylose and core  $\alpha$ 1,3-linked fucose residues attached.

Example 9: Production of Anti-CD20 and Anti-HER2 Monoclonal Antibody Having Increased ADCC Activity

IDEC-C2B8 (IDEC Pharmaceuticals Corp., San Diego, California; commercially available under the tradename Rituxan®, also referred to as rituximab; see U.S. Patent No. 5,736,137, herein incorporated by reference) is a chimeric anti-

CD20 monoclonal antibody containing human IgG1 and kappa constant regions with murine variable regions isolated from a murine anti-CD20 monoclonal antibody, IDEC-2B8 (Reff *et al.* (1994) *Blood* 83:435-445). Rituximab is licensed for treatment of relapsed B cell low-grade or follicular non-Hodgkin's lymphoma (NHL). The anti-  
5 CD20 antibody marketed as rituximab (Rituxan®) is recombinantly produced in CHO cells. The glycosylation pattern of this CHO-expressed anti-CD20 antibody reveals a heterogeneous mixture of glycoforms.

A humanized anti-ERBB2 antibody is commercially available under the tradename Herceptin® (Genentech, Inc., San Francisco, California) (see U.S. Patent  
10 No. 6,165,464, herein incorporated by reference). This recombinant humanized monoclonal antibody has high affinity for p185HER2. Early clinical trials with patients having extensive metastatic breast carcinomas demonstrate the ability of this monoclonal antibody to inhibit growth of breast cancer cells that overexpress HER2 (Baselga *et al.* (1996) *J. Clin. Oncol.* 14(3):737-744).

15 A rituximab-sequence antibody and Herceptin® anti-ERBB2-sequence antibody are recombinantly produced in *Lemna* having a wild-type glycosylation pattern, using the mAbI01 construct described above, and in *Lemna* that is genetically modified to suppress expression of both FucT and XylT, using the mAbI04 chimeric RNAi construct, with the rituximab or anti-ERBB2 heavy and light chain coding  
20 sequences replacing those for the mAbI in each of these constructs. The sequences encoding the heavy and light chains are optionally both codon-optimized with *Lemna*-preferred codons. The secreted rituximab-sequence mAb (i.e., having the amino acid sequence of the rituximab antibody) and anti-ERBB2-sequence mAb (i.e., having the amino acid sequence of the Herceptin® anti-ERBB2 mAb) are analyzed for  
25 glycosylation pattern as described above in Example 1.

The glycosylation profile for intact rituximab-sequence mAb or anti-ERBB2-sequence mAb produced in the wild-type *Lemna* comprising the mAbI01-like construct shows a heterogeneous profile with numerous peaks corresponding to multiple glycoforms. In contrast, the glycoylation profile for intact rituximab-sequence mAb or intact anti-ERBB2-sequence mAb produced in transgenic *Lemna*  
30 comprising the mAbI04-like construct shows a substantially homogeneous glycoprotein composition, with three major glycoform peaks, the largest of which corresponds to the G0 glycoform, and two very minor peaks corresponding to trace

amounts of precursor glycoforms, wherein xylose and fucose residues are not attached.

5 The rituximab-sequence and anti-ERBB2-sequence monoclonal antibody compositions having a glycosylation profile that is substantially homogeneous for the G0 glycoform are tested for ADCC activity and Fcγ receptor IIIa (Fcγ RIIIa) binding on freshly-isolated human NK cells. The rituximab-sequence and anti-ERBB2-sequence mAbs produced from *L. minor* lines engineered with the mAbI04-like construct exhibit the expected enhanced binding in view of the lack of any fucose residues, relative to the binding observed for the rituximab-sequence and anti-  
10 ERBB2-sequence mAbs produced from *L. minor* lines having the wild-type glycosylation machinery (i.e., no silencing of FucT or XylT). Furthermore, binding affinity is at least strong as for the corresponding mAb produced in CHO cell lines.

ADCC activity of the substantially homogeneous G0 glycoform of rituximab-sequence mAb and of anti-ERBB2-sequence mAb produced in the *L. minor* line  
15 engineered with the mAbI04-like construct is assayed using purified human peripheral blood mononuclear cells as effector cells (see, for example, Shinkawa *et al.* (2003) *J. Biol. Chem.* 278(5):3466-3473; herein incorporated by reference in its entirety). Activity of the G0 glycoform rituximab-sequence and anti-ERBB2-sequence mAb compositions is improved 50-1000 fold over that exhibited by the respective  
20 rituximab-sequence mAb or anti-ERBB2-sequence mAb produced in the *L. minor* line having the wild-type glycosylation machinery or produced in CHO cells.

CDC activity of the substantially homogeneous G0 glycoform of rituximab-sequence mAb produced in the *L. minor* line engineered with the mAbI04-like RNAi construct is assayed using standard assays known in the art (see, for example, the  
25 complement activation assays described in U.S. Patent Application Publication No. 2004/0167319, herein incorporated by reference in its entirety), and compared to that observed for rituximab. Non-relevant antibody serves as the negative control. In one such assay, CDC activity of the various antibodies against target Daudi cells is measured by assessing elevated membrane permeability using a propidium iodide (PI)  
30 exclusion assay, with serum from healthy volunteers serving as a complement source. Serum for complement lysis is prepared by drawing blood from healthy volunteers into autosep gel and clot activator vacutainer tubes (BD Biosciences, Rutherford, New

Jersey), which are held at room temperature for 30-60 minutes and then centrifuged at 3000 rpm for 5 minutes. Serum is harvested and stored at -80°C.

5 Briefly, for this CDC activity assay, Daudi cells are washed and resuspended in RPMI-1% BSA at  $1 \times 10^6$ /ml. Various concentrations of the substantially homogeneous G0 glycoform of rituximab-sequence mAb, rituximab, and negative control mAb are added to the Daudi cells and allowed to bind for 10-15 minutes at room temperature. Thereafter, serum as a source of complement is added to a final concentration of 20% (v/v) and the mixtures are incubated for 45 min at 37°C. The cells are then kept at 4°C until analysis. Each sample (150 µl) is then added to 10 µl  
10 of PI solution (10 µg/ml in PBS) in a FACS tube. The mixture is assessed immediately for cell lysis (number of PI-positive cells) by a FACScalibur flow cytometer and analysed using CellQuest pro software (BD Biosciences, Mountain View, California). At least 5000 events are collected for analysis with cell debris excluded by adjustment of the forward sideward scatter (FCS) threshold.

15 CDC activity of the substantially homogeneous G0 glycoform rituximab-sequence mAb is decreased relative to that exhibited by rituximab.

20 All publications and patent applications mentioned in the specification are indicative of the level of those skilled in the art to which this invention pertains. All publications and patent applications are herein incorporated by reference to the same extent as if each individual publication or patent application was specifically and individually indicated to be incorporated by reference.

## THAT WHICH IS CLAIMED:

1. A glycoprotein composition comprising a substantially homogeneous *N*-glycosylation profile, wherein at least 90% of the *N*-glycans species present in said profile are GlcNAc2Man3GlcNAc2 (G0), said profile comprising a trace amount of precursor *N*-glycan species, wherein said precursor *N*-glycan species is selected from the group consisting of Man3GlcNAc2, GlcNAc1Man3GlcNAc2 wherein GlcNAc1 is attached to the 1,3 mannose arm (MGn), GlcNAc1Man3GlcNAc2 wherein GlcNAc1 is attached to the 1,6 mannose arm (GnM), and any combination thereof.
2. The glycoprotein composition of claim 1, wherein said glycoprotein is an immunoglobulin.
3. The glycoprotein composition of claim 2, wherein said immunoglobulin comprises an Fc region selected from the group consisting of an IgG1, IgG2, IgG3, and IgG4 region.
4. The glycoprotein composition of claim 2, wherein said immunoglobulin is a monoclonal antibody.
5. The glycoprotein composition of claim 4, wherein said monoclonal antibody binds an antigen selected from the group consisting of CD20 and ERBB2 (HER2).
6. The glycoprotein composition of any one of claims 2 through 5, wherein said immunoglobulin exhibits increased binding affinity for an FcγRIII.
7. The glycoprotein composition of any one of claims 2 through 6, wherein said immunoglobulin exhibits increased antibody-dependent cellular cytotoxicity (ADCC) activity.

8. The glycoprotein composition of any one of claims 2 through 7, wherein said immunoglobulin exhibits decreased complement-dependent cytotoxicity (CDC) activity.

5 9. A substantially homogenous glycoprotein composition, wherein at least 90% of the glycoprotein present in the composition is represented by the G0 glycoform, said composition comprising a trace amount of precursor glycoform.

10 10. The glycoprotein composition of claim 9, wherein said glycoprotein is an immunoglobulin.

11. The glycoprotein composition of claim 10, wherein said immunoglobulin comprises an Fc region selected from the group consisting of an IgG1, IgG2, IgG3, and IgG4 region.

15 12. The glycoprotein composition of claim 10, wherein said immunoglobulin is a monoclonal antibody.

20 13. The glycoprotein composition of claim 12, wherein said monoclonal antibody binds an antigen selected from the group consisting of CD20 and ERBB2 (HER2).

25 14. The glycoprotein composition of any one of claims 10 through 13, wherein said immunoglobulin exhibits increased binding affinity for an FcγRIII.

15. The glycoprotein composition of any one of claims 10 through 14, wherein said immunoglobulin exhibits increased antibody-dependent cellular cytotoxicity (ADCC) activity.

30 16. The glycoprotein composition of any one of claims 10 through 15, wherein said immunoglobulin exhibits decreased complement-dependent cytotoxicity (CDC) activity.



17. A pharmaceutical composition comprising the glycoprotein composition of any one of claims 1 through 16.

18. A host cell comprising the glycoprotein composition of any one of claims 1 through 16.

19. The host cell of claim 18, wherein said host cell is a plant host cell.

20. The host cell of claim 19, wherein said plant is a duckweed.

21. A method for altering the *N*-glycosylation pattern of a heterologous polypeptide produced in a duckweed plant or duckweed plant cell or nodule, said method comprising:

(a) introducing into said duckweed plant or duckweed plant cell or nodule a nucleotide construct comprising a first nucleotide sequence that is capable of inhibiting expression or function of an  $\alpha$ 1,3-fucosyltransferase (FucT) in said duckweed plant or duckweed plant cell or nodule, and wherein said first nucleotide sequence is operably linked to a promoter that is functional in a plant cell, said first nucleotide sequence comprising in the 5'-to-3' orientation and operably linked:

(i) a FucT forward fragment, said FucT forward fragment comprising about 500 to about 800 contiguous nucleotides having at least 90% sequence identity to a nucleotide sequence of about 500 to about 800 contiguous nucleotides of SEQ ID NO:1 or SEQ ID NO:2;

(ii) a spacer sequence comprising about 200 to about 700 nucleotides immediately downstream of said FucT forward fragment;

(iii) and a FucT reverse fragment, said FucT reverse fragment having sufficient length and sufficient complementarity to said FucT forward fragment such that said first nucleotide sequence is transcribed as an RNA molecule capable of forming a hairpin RNA structure; and

(b) culturing said duckweed plant or duckweed plant cell or nodule under conditions suitable for expression of said heterologous polypeptide.

22. The method of claim 21, wherein said FucT reverse fragment comprises the complement of said FucT forward fragment or a sequence having at least 90% sequence identity to the complement of said FucT forward fragment.

23. The method of claim 21 or claim 22, wherein said FucT forward fragment comprises nucleotides (nt) 255-985 of SEQ ID NO:1 and said spacer sequence comprises nt 986-1444 of SEQ ID NO:1.

5

24. A method for altering the *N*-glycosylation pattern of a heterologous polypeptide produced in a duckweed plant or duckweed plant cell or nodule, said method comprising:

(a) introducing into said duckweed plant or duckweed plant cell or nodule a nucleotide construct comprising a first nucleotide sequence that is capable of inhibiting expression or function of an  $\alpha$ 1,3-fucosyltransferase (FucT) in said duckweed plant or duckweed plant cell or nodule, and wherein said first nucleotide sequence is operably linked to a promoter that is functional in a plant cell, said first nucleotide sequence comprising in the 5'-to-3' orientation and operably linked:

(i) a sense nucleotide sequence comprising at least 19 contiguous nucleotides having at least 90% sequence identity to a nucleotide sequence of at least 19 contiguous nucleotides of SEQ ID NO:1; and

(ii) an antisense nucleotide sequence comprising at least 19 contiguous nucleotides having at least 90% sequence identity to the complement of a nucleotide sequence of at least 19 contiguous nucleotides of said sense nucleotide sequence; wherein said first nucleotide sequence is transcribed as a small hairpin RNA having a base-paired stem length shorter than about 200 base pairs; and

(b) culturing said duckweed plant or duckweed plant cell or nodule under conditions suitable for expression of said heterologous polypeptide.

25

25. The method of claim 24, wherein said promoter operably linked to said first nucleotide sequence is a type 3 RNA polymerase III promoter.

26. The method of any one of claims 21 through 25, further comprising inhibiting expression or function of a  $\beta$ 1,2-xylosyltransferase (XylT) in said duckweed plant or duckweed plant cell or nodule.

30

27. The method of claim 26, wherein the expression or function of said XylT is inhibited by a method selected from the group consisting of:

(a) introducing a polynucleotide or polypeptide into said duckweed plant or duckweed plant cell or nodule, wherein said polynucleotide or polypeptide inhibits expression or function of said XylT in said a duckweed plant or duckweed plant cell or nodule;

(b) eliminating a gene in said duckweed plant or duckweed plant cell or nodule, wherein said gene encodes said XylT; and

(c) mutating a gene in said duckweed plant or duckweed plant cell or nodule, wherein said gene encodes said XylT.

28. The method of claim 27, comprising introducing into said duckweed plant or duckweed plant cell or nodule a nucleotide construct comprising a nucleotide sequence that is capable of inhibiting expression or function of said XylT in said duckweed plant or duckweed plant cell or nodule, wherein said nucleotide sequence that is capable of inhibiting expression or function of said XylT is operably linked to a promoter that is functional in a plant cell.

29. The method of claim 28, wherein said nucleotide sequence that is capable of inhibiting expression or function of said XylT comprises in the 5'-to-3' orientation and operably linked:

(a) a XylT forward fragment, said XylT forward fragment comprising about 500 to about 800 contiguous nucleotides having at least 90% sequence identity to a nucleotide sequence of about 500 to about 800 contiguous nucleotides of SEQ ID NO:4, SEQ ID NO:5, SEQ ID NO:19, or SEQ ID NO:20;

(b) a spacer sequence comprising about 200 to about 700 nucleotides immediately downstream of said XylT forward fragment;

(c) and a XylT reverse fragment, said XylT reverse fragment having sufficient length and sufficient complementarity to said XylT forward fragment such that said nucleotide sequence that is capable of inhibiting expression or function of said XylT is transcribed as an RNA molecule capable of forming a hairpin RNA structure.

30. The method of claim 29, wherein said XylT reverse fragment comprises the complement of said XylT forward fragment or a sequence having at least 90% sequence identity to the complement of said XylT forward fragment.

5 31. The method of claim 29 or claim 30, wherein:

(a) said XylT forward fragment comprises nucleotides (nt) 318-1052 of SEQ ID NO:4 and wherein said spacer sequence comprises nt 1053-1599 of SEQ ID NO:4; or

10 (b) said XylT forward fragment comprises nt 1-734 of SEQ ID NO:19 and said spacer sequence comprises nt 735-1282 of SEQ ID NO:19.

32. The method of claim 28, wherein said nucleotide sequence that is capable of inhibiting expression or function of said XylT comprises in the 5'-to-3' direction and operably linked:

15 (a) a sense nucleotide sequence comprising at least 19 contiguous nucleotides having at least 90% sequence identity to a nucleotide sequence of at least 19 contiguous nucleotides of SEQ ID NO:4 or SEQ ID NO:19; and

(b) an antisense nucleotide sequence comprising at least 19 contiguous nucleotides having at least 90% sequence identity to the complement of a nucleotide sequence of at least 19 contiguous nucleotides of said sense nucleotide sequence of preceding item (a);

20 wherein said nucleotide sequence that is capable of inhibiting expression or function of said XylT is transcribed as a small hairpin RNA having a base-paired stem length shorter than about 200 base pairs.

25

33. The method of claim 32, wherein said promoter operably linked to said nucleotide sequence that is capable of inhibiting expression or function of said XylT is a type 3 RNA polymerase III promoter.

30 34. A method for altering the *N*-glycosylation pattern of a heterologous polypeptide produced in a duckweed plant or duckweed plant cell or nodule, said method comprising:

(a) introducing into said duckweed plant or duckweed plant cell or nodule a nucleotide construct comprising a first nucleotide sequence that is capable of inhibiting expression or function of a  $\beta$ 1,2-xylosyltransferase (XylT) in said duckweed plant or duckweed plant cell or nodule, and wherein said first nucleotide sequence is operably linked to a promoter that is functional in a plant cell, said first nucleotide sequence comprising in the 5'-to-3' orientation and operably linked:

(i) a XylT forward fragment, said XylT forward fragment comprising about 500 to about 800 contiguous nucleotides having at least 90% sequence identity to a nucleotide sequence of about 500 to about 800 contiguous nucleotides of SEQ ID NO:6 or SEQ ID NO:21;

(ii) a spacer sequence comprising about 200 to about 700 nucleotides immediately downstream of said XylT forward fragment;

(iii) and a XylT reverse fragment, said XylT reverse fragment having sufficient length and sufficient complementarity to said XylT forward fragment such that said first nucleotide sequence is transcribed as an RNA molecule capable of forming a hairpin RNA structure; and

(b) culturing said duckweed plant or duckweed plant cell or nodule under conditions suitable for expression of said heterologous polypeptide.

35. The method of claim 34, wherein said XylT reverse fragment comprises the complement of said XylT forward fragment or a sequence having at least 90% sequence identity to the complement of said XylT forward fragment.

36. The method of claim 34 or claim 35, wherein:

(a) said XylT forward fragment comprises nucleotides (nt) 318-1052 of SEQ ID NO:4 and wherein said spacer sequence comprises nt 1053-1599 of SEQ ID NO:4; or

(b) said XylT forward fragment comprises nt 1-734 of SEQ ID NO:19 and wherein said spacer sequence comprises nt 735-1282 of SEQ ID NO:19.

37. A method for altering the *N*-glycosylation pattern of a heterologous polypeptide produced in a duckweed plant or duckweed plant cell or nodule, said method comprising:

5 (a) introducing into said duckweed plant or duckweed plant cell or nodule a nucleotide construct comprising a first nucleotide sequence that is capable of inhibiting expression or function of a  $\beta$ 1,2-xylosyltransferase (XylT) in said duckweed plant or duckweed plant cell or nodule, and wherein said first nucleotide sequence is operably linked to a promoter that is functional in a plant cell, said first nucleotide sequence comprising in the 5'-to-3' orientation and operably linked:

(i) a sense nucleotide sequence comprising at least 19 contiguous nucleotides having at least 90% sequence identity to a nucleotide sequence of at least 19 contiguous nucleotides of SEQ ID NO:4 or SEQ ID NO:19; and

10 (ii) an antisense nucleotide sequence comprising at least 19 contiguous nucleotides having at least 90% sequence identity to the complement of a nucleotide sequence of at least 19 contiguous nucleotides of said sense nucleotide sequence of preceding item (a); wherein said nucleotide sequence that is capable of inhibiting expression or function of said XylT is transcribed as a small hairpin RNA  
15 having a base-paired stem length shorter than about 200 base pairs; and

(b) culturing said duckweed plant or duckweed plant cell or nodule under conditions suitable for expression of said heterologous polypeptide.

20 38. The method of claim 37, wherein said promoter operably linked to said first nucleotide sequence is a type 3 RNA polymerase III promoter.

25 39. The method of anyone of claims 34 through 38, further comprising inhibiting expression or function of an  $\alpha$ 1,3-fucosyltransferase (FucT) in said duckweed plant or duckweed plant cell or nodule.

40. The method of claim 39, wherein the expression or function of said FucT is inhibited by a method selected from the group consisting of:

30 (a) introducing a polynucleotide or polypeptide into said duckweed plant or duckweed plant cell or nodule, wherein said polynucleotide or polypeptide inhibits expression or function of said FucT in said a duckweed plant or duckweed plant cell or nodule;

(b) eliminating a gene in said duckweed plant or duckweed plant cell or nodule, wherein said gene encodes said FucT; and



(c) mutating a gene in said duckweed plant or duckweed plant cell or nodule, wherein said gene encodes said FucT.

5 41. A method for altering the *N*-glycosylation pattern of a heterologous polypeptide produced in a duckweed plant or duckweed plant cell or nodule, said method comprising:

(a) introducing into said duckweed plant or duckweed plant cell or nodule a nucleotide construct comprising a fusion polynucleotide that is capable of inhibiting expression or function of an  $\alpha$ 1,3-fucosyltransferase (FucT) and a  $\beta$ 1,2-xylosyltransferase (XylT) in said duckweed plant or duckweed plant cell or nodule, wherein said fusion polynucleotide is operably linked to a promoter that is functional in a plant cell, said fusion polynucleotide comprising in the 5'-to-3' orientation and operably linked:

15 (i) a chimeric forward fragment, said chimeric forward fragment comprising:

(a) a first fragment comprising about 500 to about 650 contiguous nucleotides having at least 90% sequence identity to a nucleotide sequence of about 500 to about 650 contiguous nucleotides of a polynucleotide encoding said FucT; and

20 (b) a second fragment comprising about 500 to about 650 contiguous nucleotides having at least 90% sequence identity to a nucleotide sequence of about 500 to about 650 contiguous nucleotides of a polynucleotide encoding said XylT;

25 (ii) a spacer sequence comprising about 200 to about 700 nucleotides immediately downstream of said second fragment of said chimeric forward fragment;

30 (iii) and a reverse fragment, said reverse fragment having sufficient length and sufficient complementarity to said chimeric forward fragment such that said fusion polynucleotide is transcribed as an RNA molecule capable of forming a hairpin RNA structure; and

(b) culturing said duckweed plant or duckweed plant cell or nodule under conditions suitable for expression of said heterologous polypeptide.

42. The method of claim 41, wherein said first fragment comprises about 500 to about 650 contiguous nucleotides having at least 90% sequence identity to a nucleotide sequence of about 500 to about 650 contiguous nucleotides of SEQ ID NO:1; and said second fragment comprises about 500 to about 650 contiguous  
5 nucleotides having at least 90% sequence identity to a nucleotide sequence of about 500 to about 650 contiguous nucleotides of SEQ ID NO:4 or SEQ ID NO:19.

43. The method of claim 42, wherein said reverse fragment comprises the complement of said chimeric forward fragment or a sequence having at least 90%  
10 sequence identity to the complement of said chimeric forward fragment.

44. The method of claim 42 or claim 43, wherein:  
(a) said chimeric forward fragment comprises nucleotides (nt) 254-855 of SEQ ID NO:1 and nt 318-943 of SEQ ID NO:4, and wherein said spacer  
15 sequence comprises nt 944-1443 of SEQ ID NO:4; or  
(b) said chimeric forward fragment comprises nucleotides (nt) 254-855 of SEQ ID NO:1 and nt 1-626 of SEQ ID NO:19, and wherein said spacer sequence comprises nt 627-1126 of SEQ ID NO:19.

45. A method for altering the *N*-glycosylation pattern of a heterologous polypeptide produced in a duckweed plant or duckweed plant cell or nodule, said method comprising:  
(a) introducing into said duckweed plant or duckweed plant cell or nodule a nucleotide construct comprising a fusion polynucleotide that is capable of  
25 inhibiting expression or function of an  $\alpha$ 1,3-fucosyltransferase (FucT) and a  $\beta$ 1,2-xylosyltransferase (XylT) in said duckweed plant or duckweed plant cell or nodule, wherein said fusion polynucleotide is operably linked to a promoter that is functional in a plant cell, said fusion polynucleotide comprises in the 5'-to-3' orientation and operably linked:  
(i) a chimeric forward fragment, said chimeric forward  
30 fragment comprising:  
(a) a first fragment comprising about 500 to about 650 contiguous nucleotides having at least 90% sequence identity to a nucleotide

sequence of about 500 to about 650 contiguous nucleotides of a polynucleotide encoding said XylT; and

- (b) a second fragment comprising about 500 to about 650 contiguous nucleotides having at least 90% sequence identity to a nucleotide sequence of about 500 to about 650 contiguous nucleotides of a polynucleotide encoding said FucT;
- (ii) a spacer sequence comprising about 200 to about 700 nucleotides immediately downstream of said second fragment of said chimeric forward fragment;
- (iii) and a reverse fragment, said reverse fragment having sufficient length and sufficient complementarity to said chimeric forward fragment such that said fusion polynucleotide is transcribed as an RNA molecule capable of forming a hairpin RNA structure; and
- (b) culturing said duckweed plant or duckweed plant cell or nodule under conditions suitable for expression of said heterologous polypeptide.

46. The method of claim 45, wherein said reverse fragment comprises the complement of said chimeric forward fragment or a sequence having at least 90% sequence identity to the complement of said chimeric forward fragment.

47. The method of claim 45, wherein said first fragment comprises about 500 to about 650 contiguous nucleotides having at least 90% sequence identity to a nucleotide sequence of about 500 to about 650 contiguous nucleotides of SEQ ID NO:4 or SEQ ID NO:19; and said second fragment comprises about 500 to about 650 contiguous nucleotides having at least 90% sequence identity to a nucleotide sequence of about 500 to about 650 contiguous nucleotides of SEQ ID NO:1.

48. The method of claim 47, wherein said reverse fragment comprises the complement of said chimeric forward fragment or a sequence having at least 90% sequence identity to the complement of said chimeric forward fragment.

49. A nucleotide construct comprising a first nucleotide sequence that is capable of inhibiting expression or function of an  $\alpha$ 1,3-fucosyltransferase (FucT) in a

plant, and a second nucleotide sequence that is capable of inhibiting expression or function of a  $\beta$ 1,2-xylosyltransferase (XylT) in a plant, wherein said first nucleotide sequence is operably linked to a first promoter that is functional in a plant cell, and wherein said second nucleotide sequence is operably linked to a second promoter that is functional in a plant cell, wherein said first nucleotide sequence comprises in the 5'-to-3' orientation and operably linked:

(a) a FucT forward fragment, said FucT forward fragment comprising about 500 to about 800 contiguous nucleotides having at least 90% sequence identity to a nucleotide sequence of about 500 to about 800 contiguous nucleotides of SEQ ID NO:1 or SEQ ID NO:2;

(b) a spacer sequence comprising a nucleotide fragment of the sequence set forth in SEQ ID NO:1;

(c) and a FucT reverse fragment, said FucT reverse fragment having sufficient length and sufficient complementarity to said FucT forward fragment such that said first nucleotide sequence is transcribed as an RNA molecule capable of forming a hairpin RNA structure.

50. The nucleotide construct of claim 49, wherein said FucT reverse fragment comprises the complement of said FucT forward fragment or a sequence having at least 90% sequence identity to the complement of said FucT forward fragment.

51. The nucleotide construct of claim 49 or claim 50, wherein said FucT forward fragment comprises nucleotides (nt) 255-985 of SEQ ID NO:1.

52. The nucleotide construct of any one of claims 49 through 51, wherein said nucleotide fragment of the sequence set forth in SEQ ID NO:1 comprises about 200 to 700 nucleotides immediately downstream of said FucT forward fragment.

53. The nucleotide construct of claim 52, wherein said spacer sequence comprises (nt) 986-1444 of SEQ ID NO:1.

54. A nucleotide construct comprising a first nucleotide sequence that is capable of inhibiting expression or function of an  $\alpha$ 1,3-fucosyltransferase (FucT) in a plant, and a second nucleotide sequence that is capable of inhibiting expression or function of a  $\beta$ 1,2-xylosyltransferase (XylT) in a plant, wherein said first nucleotide sequence is operably linked to a first promoter that is functional in a plant cell, and wherein said second nucleotide sequence is operably linked to a second promoter that is functional in a plant cell, wherein said first nucleotide sequence comprises in the 5'-to-3' direction and operably linked:

(a) a sense nucleotide sequence comprising at least 19 contiguous nucleotides having at least 90% sequence identity to a nucleotide sequence of at least 19 contiguous nucleotides of SEQ ID NO:1; and

(b) an antisense nucleotide sequence comprising at least 19 contiguous nucleotides having at least 90% sequence identity to the complement of a nucleotide sequence of at least 19 contiguous nucleotides of said sense nucleotide sequence;

wherein said first nucleotide sequence is transcribed as a small hairpin RNA having a base-paired stem length shorter than about 200 base pairs.

55. The nucleotide construct of claim 54, wherein said promoter operably linked to said first nucleotide sequence is a type 3 RNA polymerase III promoter.

56. The nucleotide construct of any one of claims 49 through 55, wherein said XylT is a polypeptide comprising an amino acid sequence selected from the group consisting of:

(a) the amino acid sequence set forth in SEQ ID NO:6 or SEQ ID NO:21; and

(b) an amino acid sequence having at least 90% sequence identity to the amino acid sequence set forth in SEQ ID NO:6 or SEQ ID NO:21, wherein said polypeptide has XylT activity.

57. The nucleotide construct of claim 56, wherein said second nucleotide sequence comprises a sequence selected from the group consisting of:

(a) the nucleotide sequence set forth in SEQ ID NO:4 or SEQ ID NO:19, or a complement thereof;

(b) the nucleotide sequence set forth in SEQ ID NO:5 or SEQ ID NO:20, or a complement thereof;

5 (c) a nucleotide sequence having at least 90% sequence identity to the sequence of preceding item (a) or (b); and

(d) a fragment of the nucleotide sequence of any one of preceding items (a) through (c), wherein said fragment comprises at least 75 contiguous nucleotides of said nucleotide sequence.

10

58. The nucleotide construct of claim 56, wherein said second nucleotide sequence comprises in the 5'-to-3' orientation and operably linked:

15 (a) a XylT forward fragment, said XylT forward fragment comprising about 500 to about 800 contiguous nucleotides having at least 90% sequence identity to a nucleotide sequence of about 500 to about 800 contiguous nucleotides of SEQ ID NO:4, SEQ ID NO:5, SEQ ID NO:19, or SEQ ID NO:20;

(b) a spacer sequence comprising about 200 to about 700 nucleotides;

20 (c) and a XylT reverse fragment, said XylT reverse fragment having sufficient length and sufficient complementarity to said XylT forward fragment such that said second nucleotide sequence is transcribed as an RNA molecule capable of forming a hairpin RNA structure.

25 59. The nucleotide construct of claim 58, wherein said XylT reverse fragment comprises the complement of said XylT forward fragment or a sequence having at least 90% sequence identity to the complement of said XylT forward fragment.

30 60. The nucleotide construct of claim 58 or claim 59 wherein said XylT forward fragment comprises nucleotides (nt) 318-1052 of SEQ ID NO:4 or nt 1-734 of SEQ ID NO:19.



61. The nucleotide construct of any one of claims 58 through 60, wherein said spacer sequence within said second nucleotide sequence comprises a nucleotide fragment of the sequence set forth in SEQ ID NO:4 or SEQ ID NO:19.

5           62. The nucleotide construct of claim 61, wherein said nucleotide fragment of the sequence set forth in SEQ ID NO:4 or SEQ ID NO:19 comprises about 200 to 700 nucleotides immediately downstream of said XylT forward fragment.

          63. The nucleotide construct of claim 62, wherein:  
10           (a) said XylT forward fragment comprises nt 318-1052 of SEQ ID NO:4 and said spacer sequence comprises nt 1053-1599 of SEQ ID NO:4; or  
          (b) said XylT forward fragment comprises nt 1-734 of SEQ ID NO:19 and said spacer sequence comprises nt 735-1282 of SEQ ID NO:19.

15           64. The nucleotide construct of any one of claims 58 through 60, wherein said spacer sequence within said second nucleotide sequence comprises an intron.

          65. The nucleotide construct of claim 56, wherein said second nucleotide sequence comprises in the 5'-to-3' direction and operably linked:  
20           (a) a sense nucleotide sequence comprising at least 19 contiguous nucleotides having at least 90% sequence identity to a nucleotide sequence of at least 19 contiguous nucleotides of SEQ ID NO:4 or SEQ ID NO:19; and  
          (b) an antisense nucleotide sequence comprising at least 19  
contiguous nucleotides having at least 90% sequence identity to the complement of a  
25           nucleotide sequence of at least 19 contiguous nucleotides of said sense nucleotide sequence of preceding item (a);  
wherein said second nucleotide sequence is transcribed as a small hairpin RNA having a base-paired stem length shorter than about 200 base pairs.

30           66. The nucleotide construct of claim 65, wherein said promoter operably linked to said second nucleotide sequence is a type 3 RNA polymerase III promoter.

67. A nucleotide construct comprising a first nucleotide sequence that is capable of inhibiting expression or function of an  $\alpha$ 1,3-fucosyltransferase (FucT) in a plant, and a second nucleotide sequence that is capable of inhibiting expression or function of a  $\beta$ 1,2-xylosyltransferase (XylT) in a plant, wherein said first nucleotide sequence is operably linked to a first promoter that is functional in a plant cell, and wherein said second nucleotide sequence is operably linked to a second promoter that is functional in a plant cell, wherein said second nucleotide sequence comprises in the 5'-to-3' orientation and operably linked:

- (a) a XylT forward fragment, said XylT forward fragment comprising about 500 to about 800 contiguous nucleotides having at least 90% sequence identity to a nucleotide sequence of about 500 to about 800 contiguous nucleotides of SEQ ID NO:4, SEQ ID NO:5, SEQ ID NO:19, or SEQ ID NO:20;
- (b) a spacer sequence comprising a nucleotide fragment of the sequence set forth in SEQ ID NO:4 or SEQ ID NO:19;
- (c) and a XylT reverse fragment, said XylT reverse fragment having sufficient length and sufficient complementarity to said XylT forward fragment such that said second nucleotide sequence is transcribed as an RNA molecule capable of forming a hairpin RNA structure.

68. The nucleotide construct of claim 67, wherein said XylT reverse fragment comprises the complement of said XylT forward fragment or a sequence having at least 90% sequence identity to the complement of said XylT forward fragment.

69. The nucleotide construct of claim 67 or claim 68, wherein said XylT forward fragment comprises nucleotides (nt) 318-1052 of SEQ ID NO:4 or nt 1-734 of SEQ ID NO:19.

70. The nucleotide construct of any one of claims 67 through 69, wherein said nucleotide fragment of the sequence set forth in SEQ ID NO:4 or SEQ ID NO:19 comprises about 200 to 700 nucleotides immediately downstream of said XylT forward fragment.

71. The nucleotide construct of claim 70 wherein:

- (a) said XylT forward fragment comprises nt 318-1052 of SEQ ID NO:4 and said spacer sequence comprises nt 1053-1599 of SEQ ID NO:4; or
- 5 (b) said XylT forward fragment comprises nt 1-734 of SEQ ID NO:19 and said spacer sequence comprises nt 735-1282 of SEQ ID NO:19.

72. A nucleotide construct comprising a first nucleotide sequence that is capable of inhibiting expression or function of an  $\alpha$ 1,3-fucosyltransferase (FucT) in a plant, and a second nucleotide sequence that is capable of inhibiting expression or function of a  $\beta$ 1,2-xylosyltransferase (XylT) in a plant, wherein said first nucleotide sequence is operably linked to a first promoter that is functional in a plant cell, and wherein said second nucleotide sequence is operably linked to a second promoter that is functional in a plant cell, wherein said second nucleotide sequence comprises in the 5'-to-3' direction and operably linked:

10

15

(a) a sense nucleotide sequence comprising at least 19 contiguous nucleotides having at least 90% sequence identity to a nucleotide sequence of at least 19 contiguous nucleotides of SEQ ID NO:4 or SEQ ID NO:19; and

(b) an antisense nucleotide sequence comprising at least 19 contiguous nucleotides having at least 90% sequence identity to the complement of a nucleotide sequence of at least 19 contiguous nucleotides of said sense nucleotide sequence of preceding item (a);

wherein said second nucleotide sequence is transcribed as a small hairpin RNA having a base-paired stem length shorter than about 200 base pairs.

20

25

73. The nucleotide construct of claim 72, wherein said promoter operably linked to said second nucleotide sequence is a type 3 RNA polymerase III promoter.

74. The nucleotide construct of any one of claims 67 through 73, wherein said FucT is a polypeptide comprising an amino acid sequence selected from the group consisting of:

30

- (a) the amino acid sequence set forth in SEQ ID NO:3; and

(b) an amino acid sequence having at least 90% sequence identity to the amino acid sequence set forth in SEQ ID NO:3, wherein said polypeptide has FucT activity.

5           75. The nucleotide construct of claim 74, wherein said first nucleotide sequence comprises a sequence selected from the group consisting of:

- (a) the nucleotide sequence set forth in SEQ ID NO:1 or a complement thereof;
- 10           (b) the nucleotide sequence set forth in SEQ ID NO:2 or a complement thereof;
- (c) a nucleotide sequence having at least 90% sequence identity to the sequence of preceding item (a) or (b); and
- (d) a fragment of the nucleotide sequence of any one of preceding items (a) through (c), wherein said fragment comprises at least 75 contiguous  
15           nucleotides of said nucleotide sequence.

76. A nucleotide construct comprising a fusion polynucleotide that is capable of inhibiting expression or function of an  $\alpha$ 1,3-fucosyltransferase (FucT) and a  $\beta$ 1,2-xylosyltransferase (XylT) in said duckweed plant or duckweed plant cell or  
20           nodule, wherein said fusion polynucleotide is operably linked to a promoter that is functional in a plant cell, said fusion polynucleotide comprising in the 5'-to-3' orientation and operably linked:

- (a) a chimeric forward fragment, said chimeric forward fragment comprising:  
25           (i) a first fragment comprising about 500 to about 650 contiguous nucleotides having at least 90% sequence identity to a nucleotide sequence of about 500 to about 650 contiguous nucleotides of a polynucleotide encoding said FucT; and  
             (ii) a second fragment comprising about 500 to about 650  
30           contiguous nucleotides having at least 90% sequence identity to a nucleotide sequence of about 500 to about 650 contiguous nucleotides of a polynucleotide encoding said XylT;

(b) a spacer sequence comprising a fragment of said polynucleotide encoding said XylT, said fragment being about 200 to about 700 nucleotides in length ;

5 (c) and a reverse fragment, said reverse fragment having sufficient length and sufficient complementarity to said chimeric forward fragment such that said fusion polynucleotide is transcribed as an RNA molecule capable of forming a hairpin RNA structure.

10 77. The nucleotide construct of claim 76, wherein said reverse fragment comprises the complement of said chimeric forward fragment or a sequence having at least 90% sequence identity to the complement of said chimeric forward fragment.

15 78. The nucleotide construct of claim 76 or claim 77, wherein said spacer sequence comprises about 200 to 700 nucleotides immediately downstream of the second fragment of said chimeric forward fragment.

20 79. The nucleotide construct of any one of claims 76 through 78, wherein said FucT is a polypeptide comprising an amino acid sequence selected from the group consisting of:

(a) the amino acid sequence set forth in SEQ ID NO:3; and  
(b) an amino acid sequence having at least 90% sequence identity to the amino acid sequence set forth in SEQ ID NO:3, wherein said polypeptide has FucT activity.

25

80. The nucleotide construct of any one of claims 76 through 79, wherein said XylT is a polypeptide comprising an amino acid sequence selected from the group consisting of:

30 (a) the amino acid sequence set forth in SEQ ID NO:6 or SEQ ID NO:21; and  
(b) an amino acid sequence having at least 90% sequence identity to the amino acid sequence set forth in SEQ ID NO:6 or SEQ ID NO:21, wherein said polypeptide has XylT activity.

81. The nucleotide construct of claim 76, wherein said first fragment comprises about 500 to about 650 contiguous nucleotides having at least 90% sequence identity to a nucleotide sequence of about 500 to about 650 contiguous nucleotides of SEQ ID NO:1; and said second fragment comprises about 500 to about 650 contiguous nucleotides having at least 90% sequence identity to a nucleotide sequence of about 500 to about 650 contiguous nucleotides of SEQ ID NO:4 or SEQ ID NO:19.

82. The nucleotide construct of claim 81, wherein said reverse fragment comprises the complement of said chimeric forward fragment or a sequence having at least 90% sequence identity to the complement of said chimeric forward fragment.

83. The nucleotide construct of claim 81 or claim 82, wherein said chimeric forward fragment comprises:

- (a) nucleotides (nt) 254-855 of SEQ ID NO:1 and nt 318-943 of SEQ ID NO:4; or
- (b) nucleotides (nt) 254-855 of SEQ ID NO:1 and nt 1-626 of SEQ ID NO:19.

84. The nucleotide construct of any one of claims 81 through 83, wherein said spacer sequence comprises a nucleotide fragment of the sequence set forth in SEQ ID NO:4 or SEQ ID NO:19.

85. The nucleotide construct of claim 84, wherein said nucleotide fragment comprises about 200 to 700 nucleotides immediately downstream of the second fragment of said chimeric forward fragment.

86. The nucleotide construct of claim 85, wherein:

- (a) said chimeric forward fragment comprises nt 254-855 of SEQ ID NO:1 and nt 318-943 of SEQ ID NO:4 and said spacer sequence comprises nt 944-1443 of SEQ ID NO:4; or



(b) said chimeric forward fragment comprises nt 254-855 of SEQ ID NO:1 and nt 1-626 of SEQ ID NO:19 and said spacer sequence comprises nt nt 627-1126 of SEQ ID NO:19.

5           87. A nucleotide construct comprising a fusion polynucleotide that is capable of inhibiting expression or function of an  $\alpha$ 1,3-fucosyltransferase (FucT) and a  $\beta$ 1,2-xylosyltransferase (XylT) in said duckweed plant or duckweed plant cell or nodule, wherein said fusion polynucleotide is operably linked to a promoter that is functional in a plant cell, said fusion polynucleotide comprises in the 5'-to-3' orientation and operably linked:

10                   (a) a chimeric forward fragment, said chimeric forward fragment comprising:

                                  (i) a first fragment comprising about 500 to about 650 contiguous nucleotides having at least 90% sequence identity to a nucleotide sequence of about 500 to about 650 contiguous nucleotides of a polynucleotide encoding said XylT; and

                                  (ii) a second fragment comprising about 500 to about 650 contiguous nucleotides having at least 90% sequence identity to a nucleotide sequence of about 500 to about 650 contiguous nucleotides of a polynucleotide encoding said FucT;

20                   (b) a spacer sequence comprising a fragment of said polynucleotide encoding said FucT, said fragment being about 200 to about 700 nucleotides in length; and

                                  (c) a reverse fragment, said reverse fragment having sufficient length and sufficient complementarity to said chimeric forward fragment such that said fusion polynucleotide is transcribed as an RNA molecule capable of forming a hairpin RNA structure.

25           88. The nucleotide construct of claim 87, wherein said reverse fragment comprises the complement of said chimeric forward fragment or a sequence having at least 90% sequence identity to the complement of said chimeric forward fragment.

89. The nucleotide construct of claim 87 or claim 88, wherein said spacer sequence comprises about 200 to 700 nucleotides immediately downstream of the second fragment of said chimeric forward fragment.

5 90. The nucleotide construct of any one of claims 87 through 89, wherein said FucT is a polypeptide comprising an amino acid sequence selected from the group consisting of:

- (a) the amino acid sequence set forth in SEQ ID NO:3; and
  - (b) an amino acid sequence having at least 90% sequence identity
- 10 to the amino acid sequence set forth in SEQ ID NO:3, wherein said polypeptide has FucT activity.

91. The nucleotide construct of any one of claims 87 through 90, wherein said XylT is a polypeptide comprising an amino acid sequence selected from the

15 group consisting of:

- (a) the amino acid sequence set forth in SEQ ID NO:6 or SEQ ID NO:21; and
  - (b) an amino acid sequence having at least 90% sequence identity
- 20 to the amino acid sequence set forth in SEQ ID NO:6 or SEQ ID NO:21, wherein said polypeptide has XylT activity.

92. The nucleotide construct of claim 87, wherein said first fragment comprises about 500 to about 650 contiguous nucleotides having at least 90% sequence identity to a nucleotide sequence of about 500 to about 650 contiguous

25 nucleotides of SEQ ID NO:4 or SEQ ID NO:19; and said second fragment comprises about 500 to about 650 contiguous nucleotides having at least 90% sequence identity to a nucleotide sequence of about 500 to about 650 contiguous nucleotides of SEQ ID NO:1.

30 93. The nucleotide construct of claim 92, wherein said reverse fragment comprises the complement of said chimeric forward fragment or a sequence having at least 90% sequence identity to the complement of said chimeric forward fragment.

94. The nucleotide construct of claim 92 or claim 93, wherein said spacer sequence comprises a nucleotide fragment of the sequence set forth in SEQ ID NO:1.

5 95. The nucleotide construct of claim 94, wherein said nucleotide fragment comprises about 200 to 700 nucleotides immediately downstream of the second fragment of said chimeric forward fragment.

10 96. A nucleotide construct comprising a first polynucleotide sequence that is capable of inhibiting expression or function of an  $\alpha$ 1,3-fucosyltransferase (FucT) in a plant, wherein said first polynucleotide sequence comprises in the 5'-to-3' orientation and operably linked:

- 15 (a) a FucT forward fragment, said FucT forward fragment comprising about 500 to about 800 contiguous nucleotides having at least 90% sequence identity to a nucleotide sequence of about 500 to about 800 contiguous nucleotides of a polynucleotide encoding said FucT;
- (b) a spacer sequence comprising a fragment of said polynucleotide encoding said FucT, said fragment being about 200 to about 700 nucleotides in length; and;
- 20 (c) and a FucT reverse fragment, said FucT reverse fragment having sufficient length and sufficient complementarity to said FucT forward fragment such that said first polynucleotide sequence is transcribed as an RNA molecule capable of forming a hairpin RNA structure;
- wherein said first polynucleotide sequence is operably linked to a promoter that is functional in a plant cell.

25

97. The nucleotide construct of claim 96, wherein said FucT reverse fragment comprises the complement of said FucT forward fragment or a sequence having at least 90% sequence identity to the complement of said FucT forward fragment.

30

98. The nucleotide construct of claim 96 or claim 97, wherein said spacer sequence within said first polynucleotide sequence comprises about 200 to 700 nucleotides immediately downstream of said FucT forward fragment.

99. The nucleotide construct of claim 96, wherein said FucT is a polypeptide comprising an amino acid sequence selected from the group consisting of:

- 5                   (a) the amino acid sequence set forth in SEQ ID NO:3; and  
                  (b) an amino acid sequence having at least 90% sequence identity to the amino acid sequence set forth in SEQ ID NO:3, wherein said polypeptide has FucT activity.

10               100. The nucleotide construct of claim 99, wherein said FucT forward fragment comprises about 500 to about 800 contiguous nucleotides having at least 90% sequence identity to a nucleotide sequence of about 500 to about 800 contiguous nucleotides of SEQ ID NO:1 or SEQ ID NO:2.

15               101. The nucleotide construct of claim 100, wherein said FucT reverse fragment comprises the complement of said FucT forward fragment or a sequence having at least 90% sequence identity to the complement of said FucT forward fragment.

20               102. The nucleotide construct of claim 100 or claim 101, wherein said FucT forward fragment comprises nucleotides (nt) 255-985 of SEQ ID NO:1.

25               103. The nucleotide construct of any one of claims 100 through 102, wherein said spacer sequence of said first polynucleotide sequence comprises a nucleotide fragment of the sequence set forth in SEQ ID NO:1.

30               104. The nucleotide construct of claim 103, wherein said nucleotide fragment of the sequence set forth in SEQ ID NO:1 comprises about 200 to 700 nucleotides immediately downstream of said FucT forward fragment.

105. The nucleotide construct of claim 104, wherein said spacer sequence comprises nt 986-1444 of SEQ ID NO:1.

106. A nucleotide construct comprising a first polynucleotide sequence that is capable of inhibiting expression or function of a  $\beta$ 1,2-xylosyltransferase (XylT) in a plant, wherein said first polynucleotide sequence comprises in the 5'-to-3' orientation and operably linked:

- 5 (a) a XylT forward fragment, said XylT forward fragment comprising about 500 to about 800 contiguous nucleotides having at least 90% sequence identity to a nucleotide sequence of about 500 to about 800 contiguous nucleotides of a polynucleotide encoding said XylT;
- 10 (b) a spacer sequence comprising a fragment of said polynucleotide encoding said XylT, said fragment being about 200 to about 700 nucleotides in length; and;
- 15 (c) and a XylT reverse fragment, said XylT reverse fragment having sufficient length and sufficient complementarity to said XylT forward fragment such that said first polynucleotide sequence is transcribed as an RNA molecule capable of forming a hairpin RNA structure;
- wherein said first polynucleotide sequence is operably linked to a promoter that is functional in a plant cell.

20 107. The nucleotide construct of claim 106, wherein said XylT reverse fragment comprises the complement of said XylT forward fragment or a sequence having at least 90% sequence identity to the complement of said XylT forward fragment.

25 108. The nucleotide construct of claim 106 or claim 107, wherein said spacer sequence within said first polynucleotide sequence comprises about 200 to 700 nucleotides immediately downstream of said XylT forward fragment.

30 109. The nucleotide construct of claim 106, wherein said XylT is a polypeptide comprising an amino acid sequence selected from the group consisting of:

- (a) the amino acid sequence set forth in SEQ ID NO:6 or SEQ ID NO:21; and

(b) an amino acid sequence having at least 90% sequence identity to the amino acid sequence set forth in SEQ ID NO:6 or SEQ ID NO:21, wherein said polypeptide has XylT activity.

5           110. The nucleotide construct of claim 109, wherein said XylT forward fragment comprises about 500 to about 800 contiguous nucleotides having at least 90% sequence identity to a nucleotide sequence of about 500 to about 800 contiguous nucleotides of SEQ ID NO:4, SEQ ID NO:5, SEQ ID NO:19, or SEQ ID NO:20.

10           111. The nucleotide construct of claim 110, wherein said XylT reverse fragment comprises the complement of said XylT forward fragment or a sequence having at least 90% sequence identity to the complement of said XylT forward fragment.

15           112. The nucleotide construct of claim 110 or claim 111, wherein said XylT forward fragment comprises nucleotides nt 318-1052 of SEQ ID NO:4 or nt 1-734 of SEQ ID NO:19.

20           113. The nucleotide construct of any one of claims 110 through 112, wherein said spacer sequence within said first polynucleotide sequence comprises a nucleotide fragment of the sequence set forth in SEQ ID NO:4 or SEQ ID NO:19.

25           114. The nucleotide construct of claim 113, wherein said nucleotide fragment of the sequence set forth in SEQ ID NO:4 or SEQ ID NO:19 comprises about 200 to 700 nucleotides immediately downstream of said XylT forward fragment.

30           115. The nucleotide construct of claim 114, wherein:  
            (a) said XylT forward fragment comprises nt 318-1052 of SEQ ID NO:4 and said spacer sequence comprises nt 1053-1599 of SEQ ID NO:4; or  
            (b) said XylT forward fragment comprises nt 1-734 of SEQ ID NO:19 and said spacer sequence comprises nt 735-1282 of SEQ ID NO:19.



116. The nucleotide construct of any one of claims 49 through 115, further comprising at least one polynucleotide encoding a polypeptide of interest, wherein said polynucleotide encoding said polypeptide of interest is operably linked to a promoter that is functional in a plant cell.

5

117. The nucleotide construct of claim 116, wherein said polypeptide of interest is a mammalian polypeptide.

118. The nucleotide construct of claim 117, wherein said mammalian polypeptide is a monoclonal antibody.

10

119. The nucleotide construct of claim 117, wherein said mammalian polypeptide is selected from the group consisting of an interferon, erythropoietin (EPO), tissue plasminogen activator (tPA), plasminogen, blood coagulation factors, granulocyte-macrophage colony stimulating factor (GM-CSF), and therapeutic immunoglobulins.

15

120. A vector comprising a nucleotide construct according to any one of claims 49 through 119.

20

121. A plant or plant cell comprising a nucleotide construct according to any one of claims 49 through 119 stably integrated into its genome.

25

122. The plant or plant cell of claim 121, wherein said plant is a monocot or said plant cell is from a monocot.

123. The plant or plant cell of claim 122, wherein said monocot is a member of the *Lemnaceae*.

30

124. The plant or plant cell of claim 123, wherein said monocot is from a genus selected from the group consisting of the genus *Spirodela*, genus *Wolffia*, genus *Wolffiella*, genus *Landoltia*, and genus *Lemna*.

125. The plant or plant cell of claim 124, wherein said monocot is a member of a species selected from the group consisting of *Lemna minor*, *Lemna miniscula*, *Lemna aequinoctialis*, and *Lemna gibba*.

5           126. The plant or plant cell of claim 121, wherein said plant is a dicot or said plant cell is from a dicot.

10           127. A method for stably transforming a higher plant to express glycoproteins having an altered *N*-glycosylation pattern, said method comprising introducing into said higher plant the nucleotide construct according to any one of claims 50 through 95.

15           128. The method of claim 127, wherein said higher plant is a plant host expressing at least one heterologous polypeptide of interest.

          129. The method of claim 128, wherein said heterologous polypeptide of interest is a mammalian polypeptide or biologically active variant thereof.

20           130. The method of claim 129, wherein said mammalian polypeptide is a monoclonal antibody.

25           131. The method of claim 129, wherein said mammalian polypeptide is selected from the group consisting of an interferon, erythropoietin (EPO), tissue plasminogen activator (tPA), plasminogen, blood coagulation factors, granulocyte-macrophage colony stimulating factor (GM-CSF), and therapeutic immunoglobulins.

30           132. A method for stably transforming a higher plant to express a heterologous polypeptide of interest having an altered *N*-glycosylation pattern, said method comprising introducing into said higher plant the nucleotide construct according to any one of claims 116 through 119.

133. The method of claim 132, wherein the *N*-glycosylation pattern is characterized by a reduction in the attachment of  $\alpha$ 1,3-fucose residues to *N*-glycans attached to said heterologous polypeptide.

5           134. The method of claim 132, wherein the *N*-glycosylation pattern is characterized by a reduction in the attachment of  $\beta$ 1,2-xylose residues to *N*-glycans attached to said heterologous polypeptide.

10           135. The method of claim 132, wherein the *N*-glycosylation pattern is characterized by a reduction in the attachment of  $\alpha$ 1,3-fucose residues and  $\beta$ 1,2-xylose residues to *N*-glycans attached to said heterologous polypeptide.

15           136. A method for producing a heterologous mammalian glycoprotein in a higher plant, wherein said heterologous mammalian glycoprotein has an a reduction in the attachment of  $\alpha$ 1,3-fucose and  $\beta$ 1,2-xylose residues to *N*-glycans of said glycoprotein when produced in said higher plant, said method comprising:

20                   (a) introducing into said plant an expression cassette comprising a sequence encoding a mammalian polypeptide that is post-translationally processed as said glycoprotein, and a polynucleotide comprising the nucleotide construct of any one of claims 50 to 79; and

                  (b) culturing said plant under conditions suitable for expression of said glycoprotein.

25           137. The method of claim 136, wherein said mammalian polypeptide is a monoclonal antibody.

30           138. The method of claim 136, wherein said mammalian polypeptide is selected from the group consisting of an interferon, erythropoietin (EPO), tissue plasminogen activator (tPA), plasminogen, blood coagulation factors, granulocyte-macrophage colony stimulating factor (GM-CSF), and therapeutic immunoglobulins.

                  139. The method of any one of claims 127 through 138, wherein said higher plant is a monocot.

140. The method of claim 139, wherein said monocot is a member of the *Lemnaceae*.

5           141. The method of claim 140, wherein said monocot is from a genus selected from the group consisting of the genus *Spirodela*, genus *Wolffia*, genus *Wolffiella*, genus *Landoltia*, and genus *Lemna*.

10           142. The method of claim 141, wherein said monocot is a member of a species selected from the group consisting of *Lemna minor*, *Lemna miniscula*, *Lemna aequinoctialis*, and *Lemna gibba*.

15           143. The method of any one of claims 127 through 138, wherein said higher plant is a dicot.

20           144. A method for reducing heterogeneity of the *N*-glycosylation profile of a glycoprotein produced in a higher plant, said method comprising introducing into said plant the nucleotide construct according to any one of claims 50 through 99, and culturing said plant under conditions suitable for expression of said glycoprotein.

25           145. The method of claim 144, wherein said glycoprotein is an endogenous glycoprotein.

            146. The method of claim 144, wherein said glycoprotein is a heterologous glycoprotein.

            147. The method of claim 146, wherein said heterologous glycoprotein is a mammalian glycoprotein.

30           148. The method of claim 147, wherein said mammalian glycoprotein is a monoclonal antibody.

            149. The method of claim 147, wherein said mammalian glycoprotein is selected from the group consisting of an interferon, erythropoietin (EPO), tissue

plasminogen activator (tPA), plasminogen, blood coagulation factors, granulocyte-macrophage colony stimulating factor (GM-CSF), and therapeutic immunoglobulins.

5           150. The method of any one of claims 144 through 149, wherein at least 90% of the *N*-glycans within said profile are GlcNAc2Man3GlcNAc2 (G0), and wherein said profile is devoid of *N*-glycan species having fucose, xylose, or both fucose and xylose attached thereto.

10           151. The method of claim 150, wherein at least 95% of the *N*-glycans within said profile are G0.

          152. The method according to any one of claims 144 through 151, wherein said higher plant is a monocot.

15           153. The method of claim 152, wherein said monocot is a member of the *Lemnaceae*.

          154. The method of claim 153, wherein said monocot is from a genus selected from the group consisting of the genus *Spirodela*, genus *Wolffia*, genus *Wolffiella*, genus *Landoltia*, and genus *Lemna*.

          155. The method of claim 154, wherein said monocot is a member of a species selected from the group consisting of *Lemna minor*, *Lemna miniscula*, *Lemna aequinoctialis*, and *Lemna gibba*.

25           156. The method according to any one of claims 153 through 155, wherein the reduced heterogeneity of said *N*-glycosylation profile is maintained with scale-up in production of said plant, wherein production scale is increased by at least 6,500-fold.

30           157. The method according to any one of claims 153 through 156, wherein the reduced heterogeneity of said *N*-glycosylation profile is maintained with continuous clonal culture of said plant.

158. The method of claim 157, wherein the reduced heterogeneity of said *N*-glycosylation profile is maintained with continuous clonal culture of said plant for at least 8 months.

5

159. The method according to any one of claims 144 through 151, wherein said higher plant is a dicot.

10

160. A method for reducing one or more adverse side effects related to complement activation with administration of an antibody, said method comprising administering said antibody in the form of a substantially homogeneous antibody composition, wherein at least 90% of said antibody present in the composition is represented by the G0 glycoform, said composition comprising a trace amount of said antibody represented by a precursor glycoform, wherein said antibody within said composition has decreased complement-dependent cytotoxicity (CDC) activity.

15

161. The method of claim 160, wherein said antibody is a monoclonal antibody.

20

162. The method of claim 161, wherein said monoclonal antibody binds CD20 antigen.

163. The method of any one of claims 160 through 162, wherein said antibody exhibits increased binding affinity for an FcγRIII.

25

164. The method of any one of claims 160 through 163, wherein said antibody exhibits increased antibody-dependent cellular cytotoxicity (ADCC) activity.

30

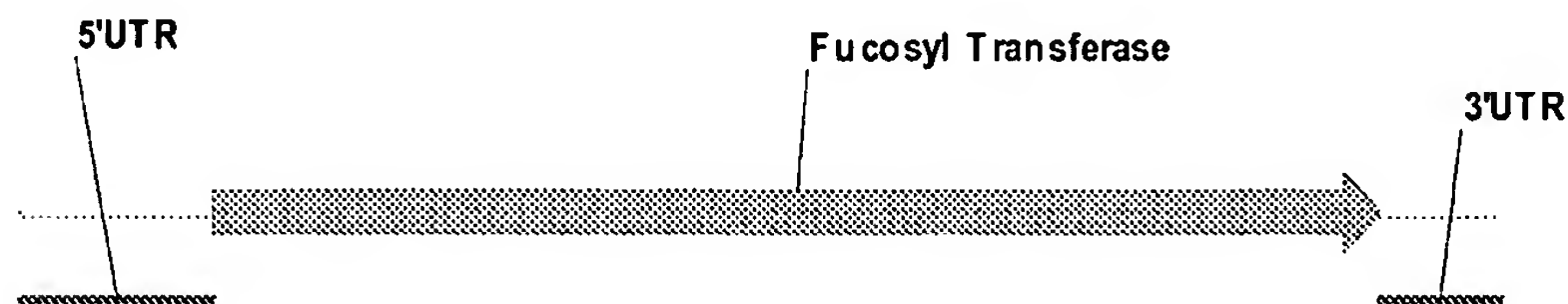
165. A composition comprising a glycoprotein produced in accordance with the method of anyone of claims 136 through 159.



1/59

## DNA and amino acid sequence information

*Lemna minor* alpha 1-3 fucosyltransferase:



L. minor RACE PCR fucosyltransferase

1865 bp

**FucT isoform #1 cDNA sequence: (UTR's in standard font, coding region in bold)**

acgcggggggaagtgggttgagtagctcagtggaattggaatgtctattagagggggaagaggggagggatccgagggg  
aagcaggaaggtgtgccgaattctcgtagatttcttcaattcctgcagatctcgtcttctctgatttcttcccagctc  
cgcccgtaggaactcaatcggaactcgatccaagttgacgaggcctacygaggaaggcgattttccgaagccctgcgatcga  
tggccacctctgctgctggtgctctcaacgcgggtggcaggggtcgggggcaggaggaggtgggtcagattgcttcccttct  
ttgtgttgatgctggtggttaggggagatctggttctcgggcggctggatgtggtcaagaacgccgctatgggtcaaaact  
ggacttcttcccacttgttttcttaccagtttcttctacacgtggtccgagaccgtcaaggaggaagaggattgcaagg  
actggctggaaagagtagatgcgggtcgattacaagagagatttccgtgtggaaccggttctggtaaatgacgctgaacagg  
attggaagttcatgttcagtgggtgtgaagttcggatcattccccggaagaacgcctgatgctacatttgggttctctcaga  
atccatcaacagtcagtggtccatcgatccatggaatcatccattatttattggagaataatcttgataatgcacgacgga  
aaggctatcaaatgtgatgacaactagttcttctgtcagatgtgcctgtcggttatttctcatgggctgaatatgatata  
tggcgctcttccagccgaaaactgctggtgcacttgcctgctgcatattatataattgaggagcacgtaatttccgcttgc  
aggcccttgatatgctcgaaaagtcgaatattaagattgattcatatggtgcttgccatcgcaaccaagacggtaaaagtgg  
acaaggtacaaactttgaagcgggtataagttcagcttagcttttgaaaactcgaacgaggatgactatgttactgagaagt  
tctttcaatctcttgcgtggtgagctattcctgttgcgtcggagccccaacattcaaaattttgagccatcttctgatt  
caattctgcacatcagggagcccaaggatgtcagttcagtcgctgagagaaatgaatttctcgttcaaatccagaagcat  
ataaccaatcactgaggtggaagtttgagggcccttctaactccttcaaaagccctgggtggacatggcagcagttcactcct  
cctgcccgttatgcattcacattgccaccaagatcagagagaaggaagagagaaaccgaatttcaagactcggccttgca  
agtgcacccgcaatgggtctacattatatacacttatacggccgcgaaagagggcacctttgacttcttatcaatcttcatga  
gatcggataatctatcactgaaagcgtggtggtcaacagttcttgagaaattcagttcttgaagcacgtgccgatttggga  
agaaggagaggccagagaggtctgaaaggaggaggagcaagctggatctttacagaatctatccagtgggcattactcagagag  
aagctctcttctcttccagttcaaacactgacaaagaacttcaaatctaccttgaatcccatccatgtgcgaagtttgaag  
tcattcttatttgatccctgaggttaattaggtcacgaattcagctaatgttgaattatgcttcaagcccacatggtatt  
tcatatcattaattgaaggcatagttagttgatattgacattttcgtctaggatcattctaaagtctatcccaatgaactt  
aa

**FucT isoform #1 predicted amino acid sequence:**

matsaagalnaggrvgrrrswvrlpffvmlvvgewflgrldvvknaamvqnwtsshlfflpvssytwsetvkeeedck  
dwlervdavykrdfvrpepvlvndaeqdwsscsvgckfgsfpgtrpdatfgfsqnpstvsvhrrsmesshyllennldnarr  
kgyqivmttsllsdvpvgyfswaeydimaplpktagalaaafisncgarnfrlqaldmleksnikidsygachrnqdgkv  
dkvqtlkrykfslafensneddyvtekffqslvagaipvvvgapniqnfpssdsilhirepkdvssvaermkflasnpa  
ynqslrwkfegpsnsfkalvdmavhsscrlicihatkirekeernpnfktrpckctrngstlyhlyarergtfdflsifm  
rsdnlskalgstvlkfsslkhpviwkkerpeslkggskldlyriypvgitqrealfsfqfntdkelqiyleshpcake  
vifi\*

**FIG. 1**

2/59

**Amino acid sequence alignment of alpha 1,3 fucosyltransferase from *Lemna minor* and other higher plants**

Total alignment positive – 89.2%

Total alignment identity – 41.1%

*Triticum aestivum*: Genbank accession# AJ582182.1*Hordeum vulgare*: Genbank accession# AJ582181.1*Oryza sativa*: Genbank accession# XM\_482907.1*Medicago sativa*: Genbank accession# AY082444.1*Arabidopsis thaliana*: Genbank accession# AY184990.1

		1	35
Lemna minor FucT	(1)	-----MASSAGALNAGGRVG	
Triticum aestivum FucT	(1)	-----MKESSESQ-----QAAQ	
Hordeum vulgare FucT	(1)	-----MESHSLN-----QAAQ	
Oryza sativa FucT	(1)	-----MKESSESQSQSQSCAG	
Medicago sativa FucT	(1)	MGLVSRT-----TTTTTQEGLEVSVVSTTV	
Arabidopsis thaliana FucT	(1)	MGVFSNLRGPRAGATHDEFPTNGSPSSSSSPSSS	
Consensus	(1)	GSHS S S A AAS	
		36	70
Lemna minor FucT	(17)	GRFSWVRLLPFEVLNLVGEIWFELGRLDNVMNAM	
Triticum aestivum FucT	(13)	RRRRCGWLLPLLVGVAFGEIAFLGRLDMAKNAAA	
Hordeum vulgare FucT	(12)	RRRRCGWLLPLLVGVAFGEIAFLGRLDMAKNAAA	
Oryza sativa FucT	(18)	RRRRCGWLLPLLVGAFFDAEIAFLGRLDMAKNAAA	
Medicago sativa FucT	(27)	PRKWSNLLPLFVAIVVIEIEIAFLGRLDMAKNAM	
Arabidopsis thaliana FucT	(36)	IKPLSNLEFLCVAEIVVIEIEISFLGRLEKVALVDE	
Consensus	(36)	RRRRCGWLLPLLVGLAVIAEIAFLGRLDMAKNAAA	
		71	105
Lemna minor FucT	(52)	VONWTSNHLFFELFVSS-----YTWSSETVK	
Triticum aestivum FucT	(48)	VESWTTSEFRLSSTWG-----ADAPP	
Hordeum vulgare FucT	(47)	VESWTTSEFRLSSTWG-----ADAPP	
Oryza sativa FucT	(53)	VESWTTSEFRLSAPARDGKAAVVVPGADAADAPP	
Medicago sativa FucT	(62)	VADLFYRSRAVEGDD-----FGL	
Arabidopsis thaliana FucT	(71)	STDFTQSPSLSQ-----APP	
Consensus	(71)	VESWTTSFYALSS G ADAPP	
		106	140
Lemna minor FucT	(76)	-----DEEDYDLERVDAVDENRDEPVA	
Triticum aestivum FucT	(69)	ESG-----DDEECERLEREDAVPYDRDFEKD	
Hordeum vulgare FucT	(68)	ESG-----DDEECERLEREDAVPYDRDFEKH	
Oryza sativa FucT	(88)	GGGEVVEEDDGGIPLCEERLEREDSVPRDPDFDKD	
Medicago sativa FucT	(81)	ETVGGDKNLELSPETCEERLEREDAVTYSNENNKD	
Arabidopsis thaliana FucT	(87)	RRSDRKKIGLFTDRSCENLMPEDSVTYSPEDETKD	
Consensus	(106)	GSG DDEECEWLEREDAVPYDRDFEKD	

**FIG. 2A**

3/59

		141	175
Lemna minor FucT	(100)	PVLVNDARCDHSSCSVGCEFGSFPGRTPDATFCF	
Triticum aestivum FucT	(97)	PVLVGGAAKDWNRCSVGCEFGFPASKTIDATFCIA	
Hordeum vulgare FucT	(96)	PVLVGGAAKDWNRCSVGCEFGFPASKTIDATFCIA	
Oryza sativa FucT	(123)	PVLVGGAAKDWNRCSVGCEFGESATKTPDATFCIA	
Medicago sativa FucT	(116)	PVFYSGADDEWKSCSVGCKTGENGDKKFAAFGNP	
Arabidopsis thaliana FucT	(122)	PVFYSGADDEWKSCSVGCKTGENGDKKFAAFGNP	
Consensus	(141)	PVLVGGAEKDWNRCSVGCEFGF ASKTPDATFGIA	
		176	210
Lemna minor FucT	(135)	QNPSTVSHRSMESSQYYLENNADNARRRGYQIVM	
Triticum aestivum FucT	(132)	QDPVESILSRSMESSQYYSENNINAARGRGYQIVM	
Hordeum vulgare FucT	(131)	QDPVESILSRSMESSQYYSENNINAARGRGYQIVM	
Oryza sativa FucT	(158)	QDPVESILSRSMESSQYYSENNINAVAPGRGYQIVM	
Medicago sativa FucT	(151)	QAGTASVLPMSMQYYSENNINAMAPRRGYHIVM	
Arabidopsis thaliana FucT	(157)	QKFGTSLSRSMESQYYPENDLAQARRRGYDIVM	
Consensus	(176)	QDPVESILSRSMESSQYYSENNIA ARRRGYQIVM	
		211	245
Lemna minor FucT	(170)	TTSLSDDVPVGYFSWAEYDIMAAPVPPKTEALAAA	
Triticum aestivum FucT	(167)	TTSLSDDVPVGYFSWAEYDIMAAPVPPKTEALAAA	
Hordeum vulgare FucT	(166)	TTSLSDDVPVGYFSWAEYDIMAAPVPPKTEALAAA	
Oryza sativa FucT	(193)	TTSLSDDVPVGYFSWAEYDIMAAPVPPKTEALAAA	
Medicago sativa FucT	(186)	TTSLSDDVPVGYFSWAEYDIMAAPVPPKTEALAAA	
Arabidopsis thaliana FucT	(192)	TTSLSDDVPVGYFSWAEYDIMAAPVPPKTEALAAA	
Consensus	(211)	TTSLSDDVPVGYFSWAEYDIMAAPVPPKTEALAAA	
		246	280
Lemna minor FucT	(205)	FISNCGARNFRLQALEMLESIDYKIDSYGSCHRN	
Triticum aestivum FucT	(202)	FISNCGARNFRLQALEMLESIDYKIDSYGSCHRN	
Hordeum vulgare FucT	(201)	FISNCGARNFRLQALEMLESIDYKIDSYGSCHRN	
Oryza sativa FucT	(228)	FISNCGARNFRLQALEMLESIDYKIDSYGSCHRN	
Medicago sativa FucT	(221)	FISNCGARNFRLQALEMLESIDYKIDSYGSCHRN	
Arabidopsis thaliana FucT	(227)	FISNCGARNFRLQALEMLESIDYKIDSYGSCHRN	
Consensus	(246)	FISNCGARNFRLQALEMLESIDYKIDSYGSCHRN	
		281	315
Lemna minor FucT	(240)	DGKVDKVTETLKRYKFSLAFENSNEEDYVTEKFFQS	
Triticum aestivum FucT	(237)	DGKVDKVTETLKRYKFSLAFENSNEEDYVTEKFFQS	
Hordeum vulgare FucT	(236)	DGKVDKVTETLKRYKFSLAFENSNEEDYVTEKFFQS	
Oryza sativa FucT	(263)	DGKVDKVTETLKRYKFSLAFENSNEEDYVTEKFFQS	
Medicago sativa FucT	(256)	DGKVDKVTETLKRYKFSLAFENSNEEDYVTEKFFQS	
Arabidopsis thaliana FucT	(262)	DGKVDKVTETLKRYKFSLAFENSNEEDYVTEKFFQS	
Consensus	(281)	DGKVDKVTETLKRYKFSLAFENSNEEDYVTEKFFQS	
		316	350
Lemna minor FucT	(275)	LVTGAI PVVVGAPNIQEFAPGEDAILHIKELDDV	
Triticum aestivum FucT	(272)	LVTGAI PVVVGAPNIQEFAPGEDAILHIKELDDV	
Hordeum vulgare FucT	(271)	LVTGAI PVVVGAPNIQEFAPGEDAILHIKELDDV	
Oryza sativa FucT	(298)	LVTGAI PVVVGAPNIQEFAPGEDAILHIKELDDV	
Medicago sativa FucT	(291)	LVTGAI PVVVGAPNIQEFAPGEDAILHIKELDDV	
Arabidopsis thaliana FucT	(297)	LVTGAI PVVVGAPNIQEFAPGEDAILHIKELDDV	
Consensus	(316)	LVTGAI PVVVGAPNIQEFAPGEDAILHIKELDDV	

FIG. 2B

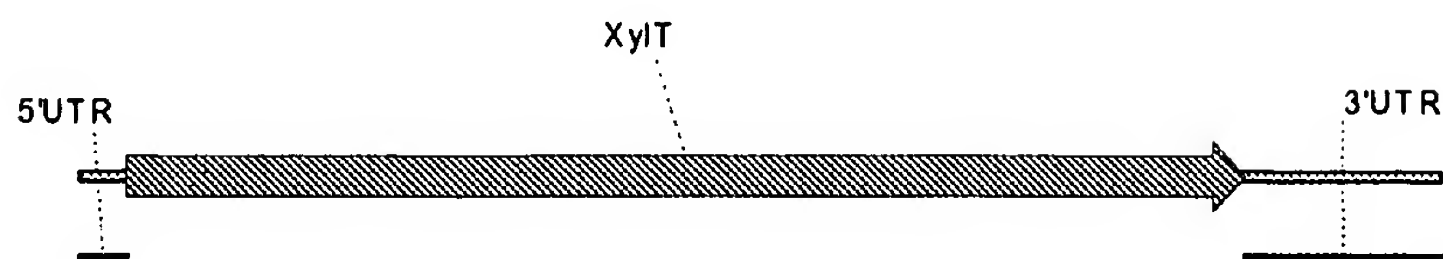


4/59

		351	385
Lemna minor FucT	(310)	SVAEKRMKFNASNEEAFNQSLRWKKEGSPNSGFKALN	
Triticum aestivum FucT	(307)	SVAKTMKHIASNEEAFNQSLRWKYDGPSSDFKALI	
Hordeum vulgare FucT	(306)	SVAKTMKHIASNEEAFNQSLRWKYDGPSSDFKALI	
Oryza sativa FucT	(333)	SVAKTMKHIASNOEAFNQSLRWKYDGPSSDFKALI	
Medicago sativa FucT	(326)	SVAKTMKHIAENEAFNHSLRWKYDGPSSDFKALI	
Arabidopsis thaliana FucT	(332)	PVAKRMKFNASNEEAFNQSLRWKYDGPSSDFKALI	
Consensus	(351)	SVAKTMKHIASNEEAFNQSLRWKYDGPSSDFKALI	
		386	420
Lemna minor FucT	(345)	DMAAVHSSCRLCIHIATKIREKEERTPKFMRNRS	
Triticum aestivum FucT	(342)	DMAAVHSSCRLCIHIATKIREKEERTPKFMRNRS	
Hordeum vulgare FucT	(341)	DMAAVHSSCRLCIHIATKIREKEERTPKFMRNRS	
Oryza sativa FucT	(368)	DMAAVHSSCRLCIHIATKIREKEERTPKFMRNRS	
Medicago sativa FucT	(361)	DMAAVHSSCRLCIHIATKIREKEERTPKFMRNRS	
Arabidopsis thaliana FucT	(367)	DMAAVHSSCRLCIHIATKIREKEERTPKFMRNRS	
Consensus	(386)	DMAAVHSSCRLCIHIATKIREKEERTPKFMRNRS	
		421	455
Lemna minor FucT	(380)	CERN-GSTVYHLFVREGRGTEFLRSQDQLTLG	
Triticum aestivum FucT	(377)	CSSK-RGTVYHLFVREGRGTEFLRSQDQLTLG	
Hordeum vulgare FucT	(376)	CSSK-RGTVYHLFVREGRGTEFLRSQDQLTLG	
Oryza sativa FucT	(403)	CSSK-RGKVYHLFVREGRGTEFLRSQDQLTLG	
Medicago sativa FucT	(396)	CERG-SETVYHLFVREGRGTEFLRSQDQLTLG	
Arabidopsis thaliana FucT	(402)	CRRGSDIVYHLFVREGRGTEFLRSQDQLTLG	
Consensus	(421)	CSSK RGTVYHLFVREGRGTEFLRSQDQLTLG	
		456	490
Lemna minor FucT	(414)	ALGSTVLEKFSLSIKHVPVWKDERPSSIKGGDELKV	
Triticum aestivum FucT	(411)	ALESVLAKEFSLSIKHVPVWKDERPSSIKGGDELKV	
Hordeum vulgare FucT	(410)	ALGSTVLEKFSLSIKHVPVWKDERPSSIKGGDELKV	
Oryza sativa FucT	(437)	ALESVLAKEFSLSIKHVPVWKDERPSSIKGGDELKV	
Medicago sativa FucT	(430)	SEKSAVLEKFSLSIKHVPVWKDERPSSIKGGDELKV	
Arabidopsis thaliana FucT	(437)	ALESVLAKEFSLSIKHVPVWKDERPSSIKGGDELKV	
Consensus	(456)	ALESVLAKEFSLSIKHVPVWKDERPSSIKGGDELKV	
		491	525
Lemna minor FucT	(449)	YKIYPGILTORQALY FQF DDAELARYIK HPCA	
Triticum aestivum FucT	(446)	YKIYPGILTORQALY FQF DDAELARYIK HPCA	
Hordeum vulgare FucT	(445)	YKIYPGILTORQALY FQF DDAELARYIK HPCA	
Oryza sativa FucT	(472)	YKIYPGILTORQALY FQF DDAELARYIK HPCA	
Medicago sativa FucT	(465)	YKIYPGILTORQALY FQF DDAELARYIK HPCA	
Arabidopsis thaliana FucT	(472)	YKIYPGILTORQALY FQF DDAELARYIK HPCA	
Consensus	(491)	YKIYPGILTORQALY FQF DDAELARYIK HPCA	
		526	
Lemna minor FucT	(484)	KFEVIFV-	
Triticum aestivum FucT	(481)	KFEVIFV-	
Hordeum vulgare FucT	(480)	KFEVIFV-	
Oryza sativa FucT	(507)	KFEVIFV-	
Medicago sativa FucT	(500)	KFEVIFV-	
Arabidopsis thaliana FucT	(507)	KFEVIFV-	
Consensus	(526)	KFEVIFV-	

FIG. 2C

5/59

***Lemna minor* beta 1-2 xylosyltransferase:****Lemna minor beta 1-2 xylosyltransferase**

1860 bp

**XylT isoform #1 cDNA sequence: (UTR's in standard font, coding region in bold)**

ggcttccaacggaggatctcgagctgaagaatcttcatgactgaagaattcatgtgatcccatggctttggtgaactcgc  
 gagggagcagggtcagacgcacgcgaagcccaccttcgttttctctcttgatcaacgtagtctgtctcctgtacttttcc  
 gtcagaaccctaattccattcccgcagcttgtcttcacggggaatgcgacaaacccccgatttttagtgactccccggcgat  
 ggaacttgaagccatggccgattcttctctctcttctgccatgggtgccgagctcccaccctgccagggtctcctgcgaag  
 cctacttccggcaacagcttcaaccgcccggacggagatgctgaagaaggtagaggggaagaggatgggttccagtgcctgtaca  
gcgatactcttcgaagttctgtttgccaggagggaatttgccgatggacccggaaaggattaggatgtcgaaggggggg  
aagatctagaggaggtgatgaagagagaggaggaagaagaattgcccaaattcgaggaggggtcgttccagattgaatctg  
gttatggaagcggagggggaagttggagagagaattgcgactgacgaggtcctcagataatgttqtgccgaaggcgctgttc  
atgtacataccatgcgcaatctcatcagttcgattcagattgttgggtcccgggcatcttcaatgctctcagtggatcgacg  
aaccggttcttcttctgacacgcttcgaatacgcgaatctcttccacacgctcaccgactggtacagcgcttacgcaagct  
cgaggattgccaacttgcccttctcgccctcacttaattttcgtcgatggccattgcagggcggaacagttagaggacatgt  
ggagagccctgttctcgacgctccgatactccaagaacttctcccagccaatctgcttccgccacgtcgtcctctcacctc  
tgggctatgagacggctctcttcaaaggcctatcagagagcttcagctgtgagggagctccggccaatcggtcaaagtca  
accccgatgaccagaagactgcaagactggctgaattcggagagatgatcagagccgcctttgactttcctgtcgttgacc  
cgtccattgacccgttgaccaaattccatctcttctgctgcccggggaagattacgtggcgccaccacggccacagtgaggag  
tggagtgcgggctgaccaacgagcaagaggtgtttgactttctgcacaattgggcaagtcacacagagcgaggtgcaaca  
tcagtatggtcaacgggcttttcgcgcacatgggaatgaaggaacagttgaagggcattaaggaagcttcgggtggtggtgg  
ggggccacggggctgggtttgacccatctgggtggcagcaaggtcaacgacagttgttcttgagattctgagttagtcaatacc  
gtagaccgcactttcaactgatttctcgggtggaaagggttggactaccatgcaattaatcttgcggggtcatttgctgacc  
ctcgggaggtgggtcgagaaattgactggcatagttgacagggttggatgttgaaagagaagtgaaggtcaacatttgggaatt  
 ttaactttaaggggtggttaacaattgagcggcattgtcaacgggtttggatgctgggaaaagtgaataatcaacacttga  
 gttctgacattgaaggcaagacgtggaattttgatggtgttgaggatatgttgatgtggagttctgatgaattaaagcagg  
 gggtgatcatttgccagtgggaattatgttgggtgtaagagagaagggggagaataaacagtggttagagagctatgctgg

**XylT isoform #1 predicted amino acid sequence:**

malvnsrgsrvrriakptfvfllinvcllyffrqnpnpipdaclhgecdkppilvtprrrwnlkpwpilpsflpwvpsshp  
 aqgsceayfgnsfnrrtemlkkvegrgwfqclysdtlrssvcgggnlrmderirmskgedleevmkreeeeelpkfeeg  
 sfqiesgygsggevgeriatdevldnvvpkgavhvhtmrnlissiqivgpghlqcsqwidepvlvtrfeyanlfhtvtdw  
 ysayassrianlpsrphlifvdghcraeqledmwralfstvrsknfsqpicfrhvvlsplgyetalfkglsefscegap  
 anrlkvnppddqktarlaefgemiraafdfpvvdpsidpltksilfvrredyvahprhsgrrvesrltneqevfdflhnwash  
 hrgrcnismvnglfahmgmkeqlkaimeasvvvgahgagtlhlvaarsttvvleilssqyrrphfqlisrwkglldyhainl  
 agsfadprevvekltgivdrllgc\*

FIG. 3

6/59

# **Amino acid sequence alignment of beta 1,2 xylosyltransferase from *Lemna minor* and other higher plants**

Total alignment positive – 76.6%

Total alignment identity – 33.7%

*Medicago sativa*: Genbank accession# AY302251.1

*Oryza sativa*: Genbank accession# XM\_483221.1

*Arabidopsis thaliana*: Genbank accession# AJ272121.1

*Saccharum officinarum*: Genbank accession# AJ864704.1

*Nicotiana tabacum*: Genbank accession# AJ627182.1

		1	35
Lemna minor XylT	(1)	-----MALVNSRGSRVRLA	PTFVFLLS
Medicago sativa XylT	(1)	-----MNEKN	QLEFFELF
Oryza sativa XylT	(1)	MMPVRTYHHHHHHNNSNNH	RLIPRLIAVFA
Arabidopsis thaliana XylT	(1)	-----MSRNPK	LETFYMLL
Saccharum officinarum XylT	(1)	---MMAGRAHGHRN----	RLRLIPRLILVFAV
Nicotiana tabacum XylT	(1)	-----MNEKK	KFLVSLFAI
Consensus	(1)		RL RKILKILL LFAI
		36	70
Lemna minor XylT	(25)	NV	CLYRERQNFNPIPDACLHGEC DKPPILV
Medicago sativa XylT	(15)	NS	ELCNYEHHF
Oryza sativa XylT	(36)	Y	AFAYLRHOSPHPHPHAADPERDAVDAAGG
Arabidopsis thaliana XylT	(19)	NS	FLIYEFHSSFSPEQSQ---
Saccharum officinarum XylT	(28)	Y	AFAYLRHOSPHPHPHAADPERDAVDAAGG
Nicotiana tabacum XylT	(16)	NS	FLIYEFHSSFSPEQSQ---
Consensus	(36)	NSVSL	IYFL HSS P S
		71	105
Lemna minor XylT	(60)	R-----	WNLKPWPILPS
Medicago sativa XylT	(47)	SPFRRPF--	IEKHK
Oryza sativa XylT	(71)	GGGGGAVDRVRVEAFSSQ	KFPRI
Arabidopsis thaliana XylT	(51)	N-----	QSAIQ
Saccharum officinarum XylT	(63)	VG-----	APPYSH
Nicotiana tabacum XylT	(51)	ENHHDNF--	HSSVT
Consensus	(71)		S KPWPI LPSYLPWS SS
		106	140
Lemna minor XylT	(81)	PAQ--	GSCEYFGNSEN
Medicago sativa XylT	(80)	STP-L	RSCGYFGNGFTRRLVDFSSRAE
Oryza sativa XylT	(106)	RPPPKH	SCEGYFGNGFTRRLVDFSSRAE
Arabidopsis thaliana XylT	(74)	NLP-T	GSCGYFGNGFTRRLVDFSSRAE
Saccharum officinarum XylT	(88)	PP---	HSCGYFGNGFTRRLVDFSSRAE
Nicotiana tabacum XylT	(83)	NVS-L	RSCGYFGNGFTRRLVDFSSRAE
Consensus	(106)	P	SCEGYFGNGFTRRLVDFSSRAE

**FIG. 4A**



7/59

		141	175
Lemna minor XylT	(106)	----GRGTFQCLYSDELRSNCCCGNSRMDPERR	
Medicago sativa XylT	(107)	---GRGGGWFRCHYSETLRSSICEGGRVPMVVERI	
Oryza sativa XylT	(133)	---GGGGGWFRCHHGETLRSSICEGGRVPMDEGLI	
Arabidopsis thaliana XylT	(102)	---EGESSWFRCHYSETLQSSICEGRNIPMVFDRI	
Saccharum officinarum XylT	(112)	---GGGGGWFRCHHSETLGSSICEGARVDEPALI	
Nicotiana tabacum XylT	(117)	NTVSGDGGTFCHHSETLQSSICEGARINMPEEI	
Consensus	(141)	GGGGGWFRCHYSETLRSSICEGGRVMDPDRI	
		176	210
Lemna minor XylT	(136)	RMSGGEDLEVMKKEEEELPKFEESFQIESGY	
Medicago sativa XylT	(139)	KVANGGESLGVIIGFSEDEELLEDGAEVDEGV	
Oryza sativa XylT	(165)	AMSRGGEPLDQVMGRKEEEELPKFEPSALQVEA-A	
Arabidopsis thaliana XylT	(134)	VMSRGGEKLEEVHGRKEEEELPATROGAEVAEEV	
Saccharum officinarum XylT	(144)	AMSRGGEPLDQVMGRKEEEELPKFEPSALQVESPA	
Nicotiana tabacum XylT	(152)	LMSRGGEKLESVHGRSEDEELPVEKNGAFQKVTD	
Consensus	(176)	MSRGGE LEEVMGRAEEEEELPKFE GAFQVEG	
		211	245
Lemna minor XylT	(171)	SGGE-----VGEATATDEVIEDNVV	
Medicago sativa XylT	(174)	GFYGG-----KMYDEGELEDRYV	
Oryza sativa XylT	(199)	KKTGP-----LVNAGELEDAYV	
Arabidopsis thaliana XylT	(169)	SSRLGFKRHRRFGGEGGSAVSRLVNDENIENEYN	
Saccharum officinarum XylT	(179)	AGTGP-----LVNAGELEDYV	
Nicotiana tabacum XylT	(187)	KLKIG-----KLVDEKILNKYL	
Consensus	(211)	A K G	KRLVDDGFLN YV
		246	280
Lemna minor XylT	(191)	PKGANHVHMRNLESSLQVGGPHCCGQWDEPV	
Medicago sativa XylT	(192)	PRGEIMRHIMRLDIGMKVYGRKEFKCDKNEEPT	
Oryza sativa XylT	(216)	PTGGIGMHTMRSLIDSGRVVPGEDHCSQWVEEPT	
Arabidopsis thaliana XylT	(204)	QEGGEDHMTMRDLVASTRAVDTNDFVCEEWVEEPT	
Saccharum officinarum XylT	(196)	PNGEIRMTMRALLDSARVVPGEDEHCSQWVEEPT	
Nicotiana tabacum XylT	(205)	PEEATSRHTMRSLIDSLQVGADEFHTSEWVEEPS	
Consensus	(246)	P GGI RHTMRDLIDSIRVVGPGELHCSQWIEEPT	
		281	315
Lemna minor XylT	(226)	LLVTRFEYANLFHTVTDWYSAYASSRIANLPSRPH	
Medicago sativa XylT	(227)	LLVTRFEYANLFHTVTDWYSAYVSSRVTLGLPSRPH	
Oryza sativa XylT	(251)	LLVTRFEYANLFHTVTDWYSAYVSSRVTDLENRPN	
Arabidopsis thaliana XylT	(239)	LLVTRFEYANLFHTVTDWYSAYVSSRVTLGLNRP	
Saccharum officinarum XylT	(231)	LLVTRFEYANLFHTVTDWYSAYVSSRVINLEDRPN	
Nicotiana tabacum XylT	(240)	LLVTRFEYANLFHTVTDWYSAYVSSRVTLGLPSRPH	
Consensus	(281)	LLVTRFEYANLFHTVTDWYSAYVSSRVTLGLPSRPH	
		316	350
Lemna minor XylT	(261)	LVFVDGHCKAEQLEEMWALFSSVRYSSNFSQETC	
Medicago sativa XylT	(262)	LVFVDGHCKAP-LEETWKALFSSVRYAKNFSGTVG	
Oryza sativa XylT	(286)	LVFVDGHCKAQ-LEQTNWALFSSNRYVKNFSGPVC	
Arabidopsis thaliana XylT	(274)	LVFVDGHCTTQ-LEETWTALFSGRYAKNFSKPVG	
Saccharum officinarum XylT	(266)	LVFVDGHCKAP-LEQTNWALFSSNRYIKFSFGPVC	
Nicotiana tabacum XylT	(275)	LVFVDGHCETO-LEETWKALFSSVRYAKNFSGPVC	
Consensus	(316)	LVFVDGHCKAQ LEETWKALFSSVRYAKNFSGPVC	

FIG. 4B

8/59

		351		385
Lemna minor XylT	(296)	FRFVLSPLGYETALFKGLSESFSCEGAFAR	KV	
Medicago sativa XylT	(296)	FRHAILSPGLGYETALFKGLSESFCGASAGE	LR	
Oryza sativa XylT	(320)	FRHAILSPGLGYETALFKGLSESFSCEGASAES	RE	
Arabidopsis thaliana XylT	(308)	FRHAILSPGLGYETALFKGLSGETDCKGDSAHN	LWQ	
Saccharum officinarum XylT	(300)	FRFVLSPLGYETALFKGLSESFSCEGASASLRN		
Nicotiana tabacum XylT	(309)	FRHAILSPGLGYETALFKGLSESFDONGASAHDL	WQ	
Consensus	(351)	FRHAILSPGLGYETALFKGLSESISCEGASAQ	LWQ	
		386		420
Lemna minor XylT	(331)	NPDDQKTARLSEFGEMIRAAFDLPVDPSP	DPLRK	
Medicago sativa XylT	(331)	KPDNEKTARLSEFGEMIRASEGLPDLNHG	PIF	
Oryza sativa XylT	(355)	KPDHOKTARLSEFGEMILASEDLRDLSS	KTS	
Arabidopsis thaliana XylT	(343)	NPDDKTARLSEFGEMIRANISLEPNRHSLE	NPL	
Saccharum officinarum XylT	(335)	RPDHOKTARLSEFGEMIASEDLLEDGIVPS	KTS	
Nicotiana tabacum XylT	(344)	NPDDKTARLSEFGEMIRAAFDLPDRON	PPTV	
Consensus	(386)	NPDDQKTARLSEFGEMIRAAFGLPVD	I KTS	
		421		455
Lemna minor XylT	(366)	-----LNVPPEDYLAHPRHSGKVESRLS		
Medicago sativa XylT	(366)	GG-----HNALEFVRREDYLAHPRHSGKVESRLS		
Oryza sativa XylT	(389)	NG-----LNVLEVRREDYLAHPRHSGKVESRLS		
Arabidopsis thaliana XylT	(378)	SSSSSASVYNVLEVRREDYLAHPRHSGKVESRLS		
Saccharum officinarum XylT	(370)	GG-----LNVLEVRREDYLAHPRHSGKVESRLS		
Nicotiana tabacum XylT	(379)	G---P-----NVLEVRREDYLAHPRHSGKVESRLS		
Consensus	(421)	SG	NVLEVRREDYLAHPRHSGKVESRLS	
		456		490
Lemna minor XylT	(391)	NEQVEFDFLHNWASHHR--G--NENISNVNGLFAHM		
Medicago sativa XylT	(394)	NEALVEESLKSAAAN----YKGKINNVNGLFAHM		
Oryza sativa XylT	(417)	NEKEVYDATEGWAK-----GOKCKINNVNGLFAHM		
Arabidopsis thaliana XylT	(413)	NEEEVFDSLHHWVATGSTGLTFCGINNVNGLFAHM		
Saccharum officinarum XylT	(398)	NEQVEGANEKWAKK-----GOKCKINNVNGLFAHM		
Nicotiana tabacum XylT	(406)	NEEQVEDSEKSNALN----HSECKINNVNGLFAHM		
Consensus	(456)	NE EVFDSL WAA	G KCKINNVNGLFAHM	
		491		525
Lemna minor XylT	(423)	GMKEQRAITMEASVNIIGAHGAGLTHIVARSTTV		
Medicago sativa XylT	(425)	SMKEQRAIQDASVNIIGAHGAGLTHIVSALPKTV		
Oryza sativa XylT	(447)	NMKEQRAIQDASVNIIGAHGAGLTHIVSATPDTKV		
Arabidopsis thaliana XylT	(448)	SMKEQRAIQDASVNIIGAHGAGLTHIVSATPDNTV		
Saccharum officinarum XylT	(429)	SMKEQRAIQDASVNIIGAHGAGLTHIVSATPDTKV		
Nicotiana tabacum XylT	(437)	SMKEQRAIQDASVNIIGAHGAGLTHIVSAAAPKAV		
Consensus	(491)	SMKEQLRAIQDASVNIIGAHGAGLTHIVSATP	TVI	
		526		560
Lemna minor XylT	(458)	LEIISQYRPPHFALISHWKLEYHAINLAGSYAD		
Medicago sativa XylT	(460)	LEIISQYRPPHFAYTANWNGLEYHAINEDGSTAN		
Oryza sativa XylT	(482)	LEIISQYRPPHFALISHWKLEYHAINLPSSYAR		
Arabidopsis thaliana XylT	(483)	FEIISVEQYRPPHFELISHWKLEYHAINHLANSRA		
Saccharum officinarum XylT	(464)	LEIISQYRPPHFALISHWKLEYHAINLPSSYAS		
Nicotiana tabacum XylT	(472)	LEIISSEYRPPHFALISHWKLEYHPTYLEGSSYAD		
Consensus	(526)	LEIIS YRPPHFALIARWKLEYHAINL	GSYAD	

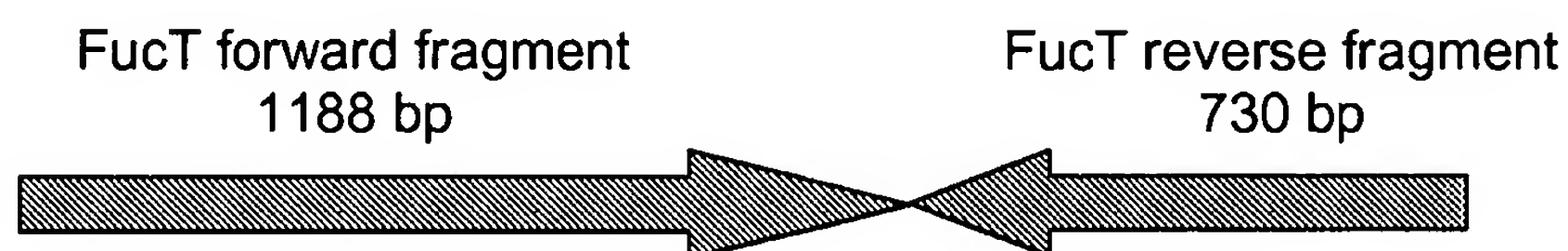
FIG. 4C

9/59

		561	578
Lemna minor XylT	(493)	PREVENFGINDRIG-	
Medicago sativa XylT	(495)	PETVINAIVSINKSLRC-	
Oryza sativa XylT	(517)	VTQVENEISNILEGFGC-	
Arabidopsis thaliana XylT	(518)	PTAVIERIEINKSLGC-	
Saccharum officinarum XylT	(499)	IPQVTSELNILEGLGC-	
Nicotiana tabacum XylT	(507)	EPVTDNISLILSLGC-	
Consensus	(561)	P DVIEKLS ILKSLGC	

FIG. 4D

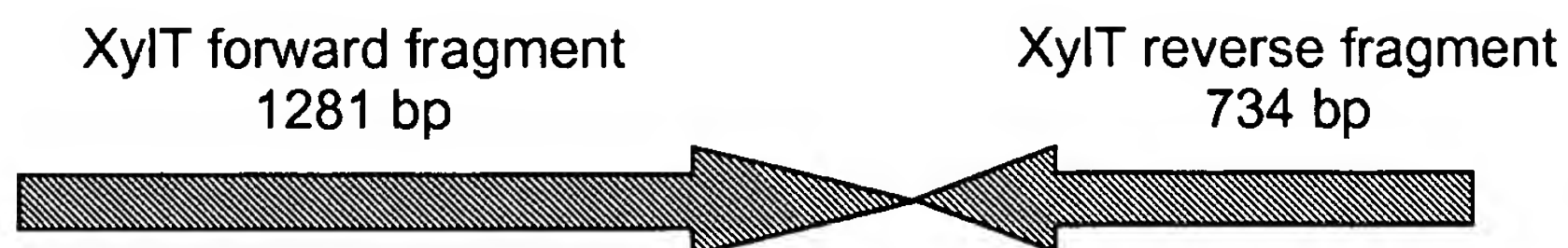
10/59

**FucT RNAi design (single gene KO):**

- Fucosyltransferase (FucT) RNAi expression cassette was designed to express a hairpin RNA with 730 bp of double-stranded RNA sequence and a 458 bp spacer sequence. The RNAi expression cassette comprises a FucT forward fragment (nt 255-985 of SEQ ID NO:1), the spacer sequence (nt 986-1444 of SEQ ID NO:1), and the FucT reverse fragment (antisense version of nt 255-985 of SEQ ID NO:1).

**FIG. 5**

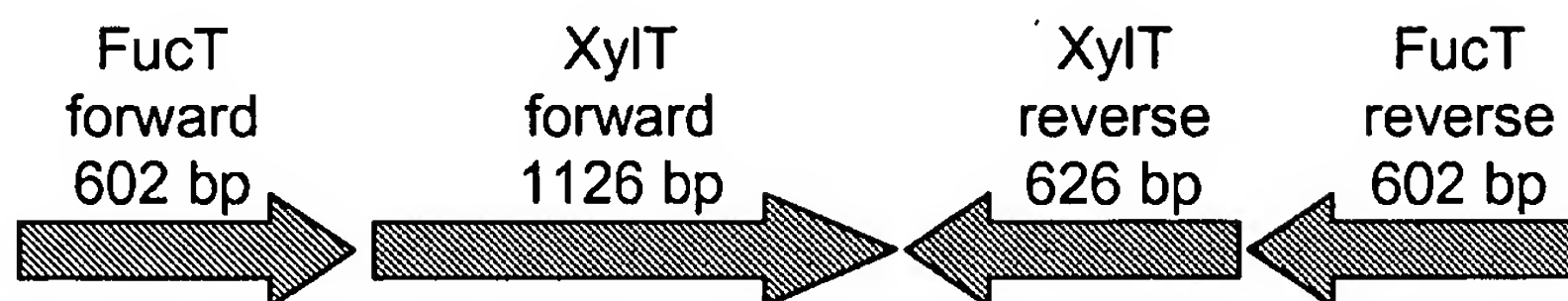
11/59

**XylT RNAi design (single gene KO):**

- Xylosyltransferase (XylT) RNAi expression cassette was designed to express a hairpin RNA with 734 bp of double-stranded RNA sequence and a 547 bp spacer sequence. The RNAi expression cassette comprises the XylT forward fragment (nt 318-1052 of SEQ ID NO:4), the spacer sequence (nt 1053-1599 of SEQ ID NO:4), and the XylT reverse fragment (antisense version of nt 318-1052 of SEQ ID NO:4).

**FIG. 6**

12/59

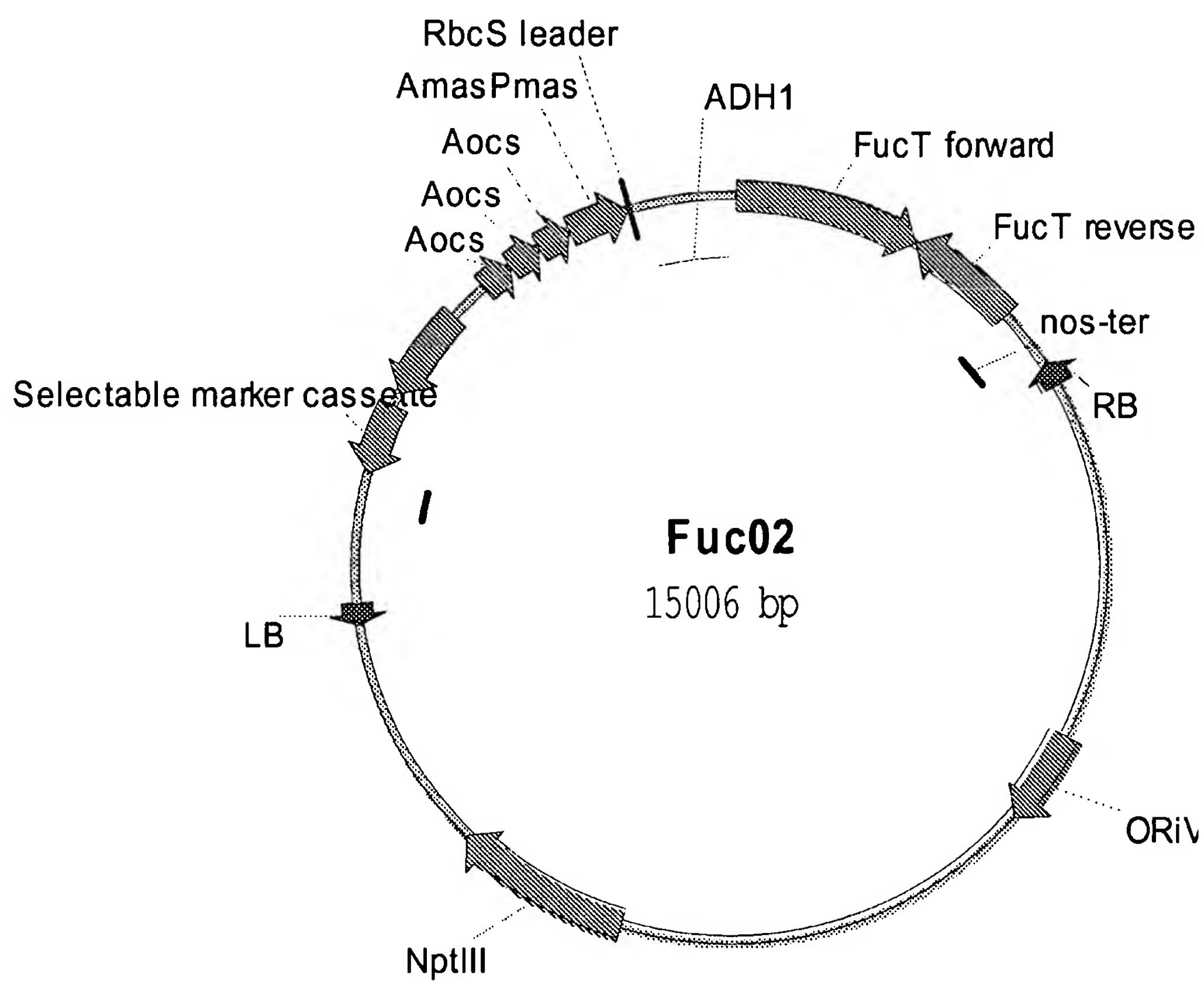
**FucT + XylT chimeric RNAi design:**

- Chimeric FucT + XylT RNAi expression cassette was designed to express a hairpin RNA with 602 bp of double-stranded FucT sequence, 626 bp of double-stranded XylT sequence, and a 500 bp spacer sequence. The RNAi expression cassette comprises in the 5' to 3' direction and operably linked: the FucT forward fragment (nt 254-855 of SEQ ID NO:1), the XylT forward fragment (nt 318-943 of SEQ ID NO:4), a spacer sequence (nt 944-1443 of SEQ ID NO:4), the XylT reverse fragment (antisense version of nt 318-943 of SEQ ID NO:4), and the FucT reverse fragment (antisense version of nt 254-855 of SEQ ID NO:1).

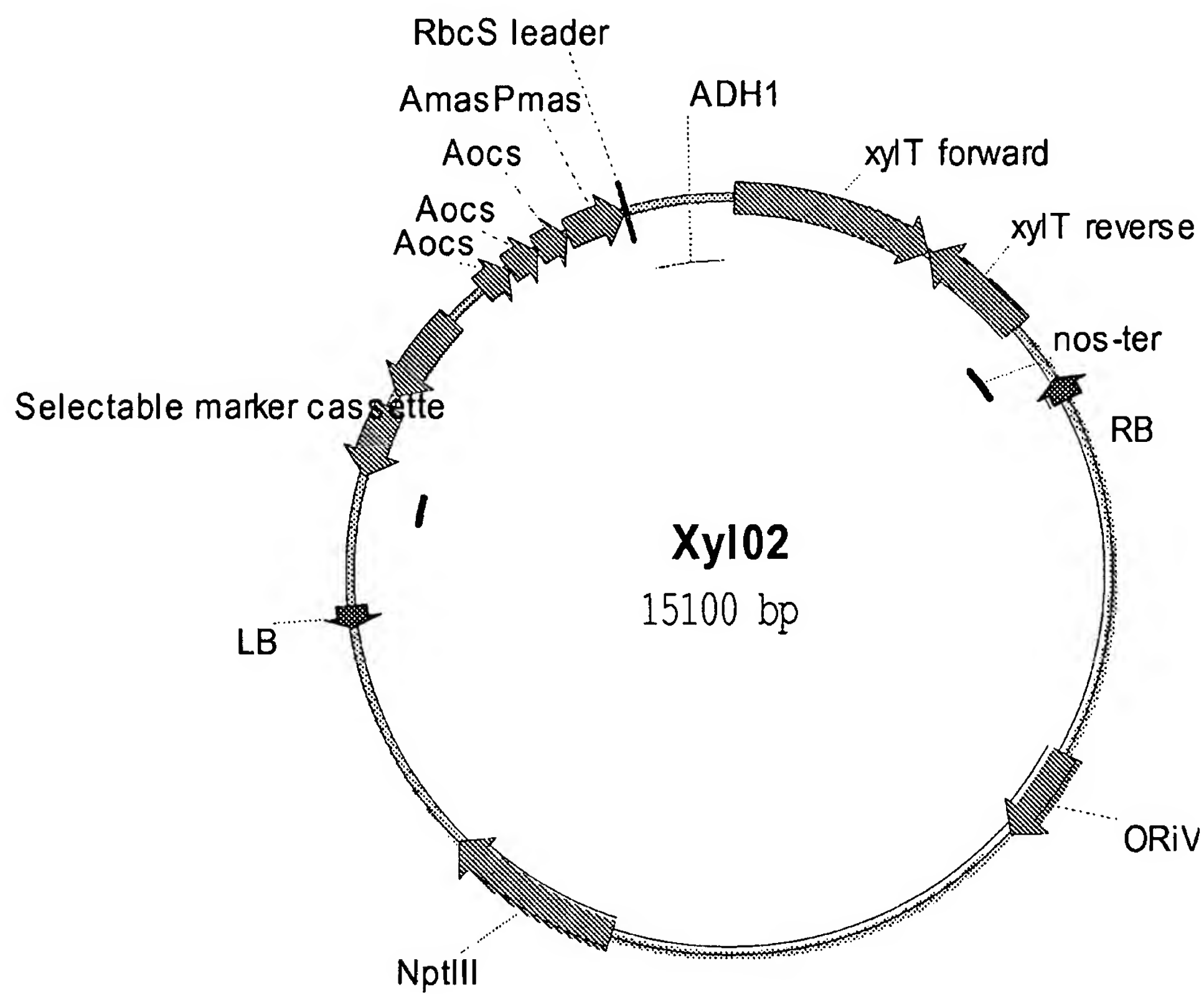
**FIG. 7**



13/59

**FucT RNAi expression cassette****FIG. 8**

14/59

**XylT RNAi expression cassette****FIG. 9**

15/59

**FucT / XylT double knockout  
Chimeric RNAi expression cassette**

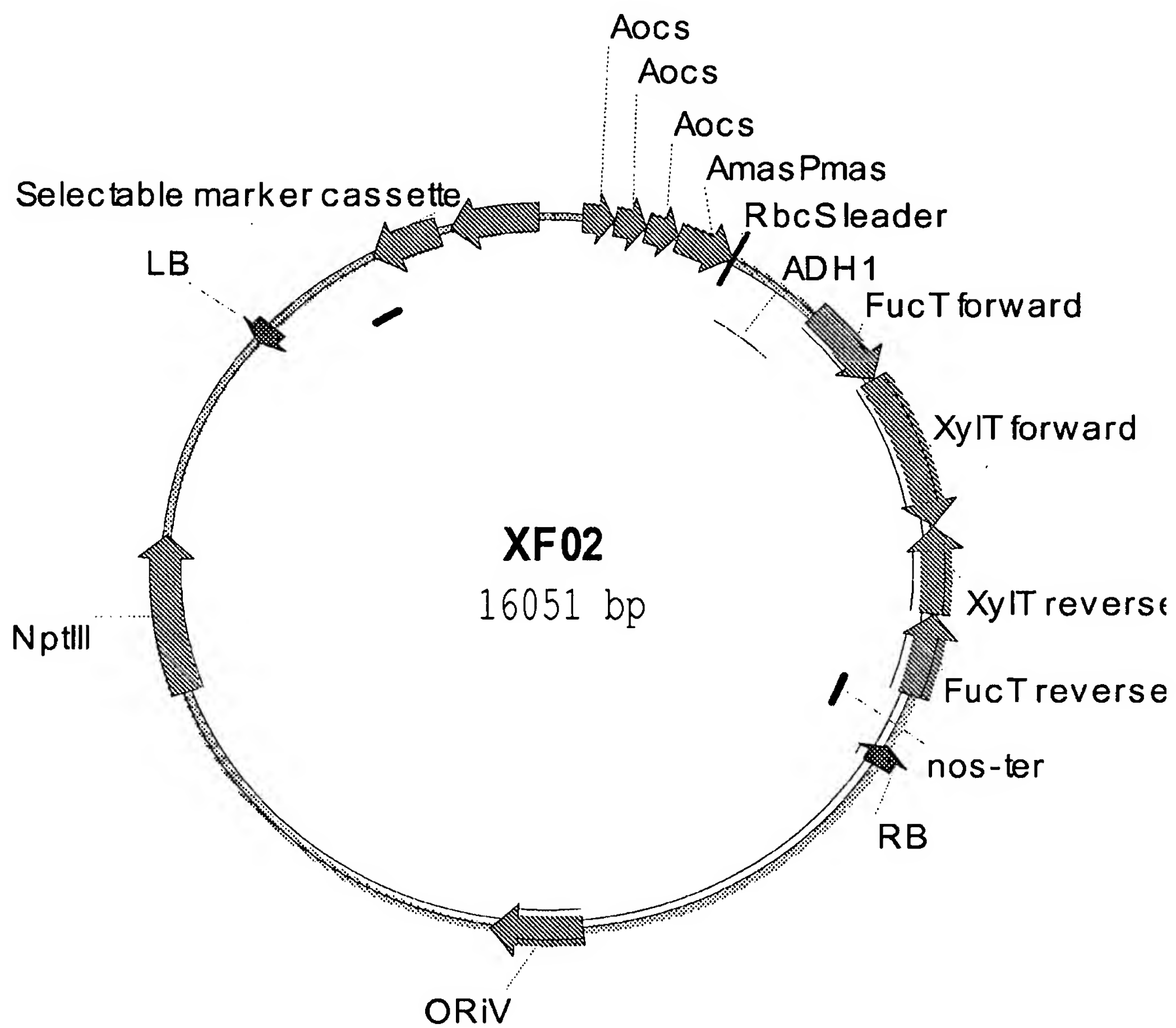
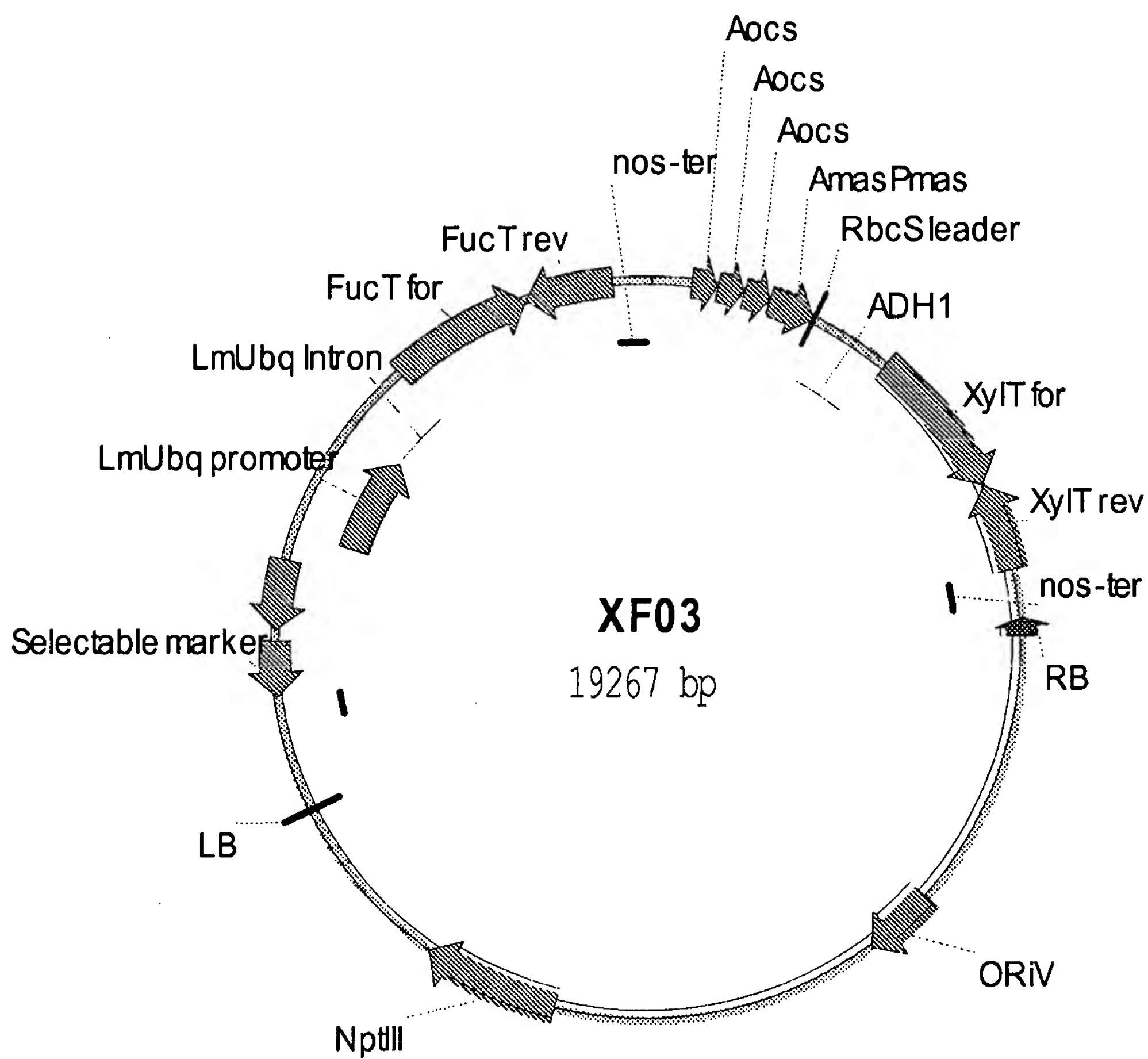


FIG. 10

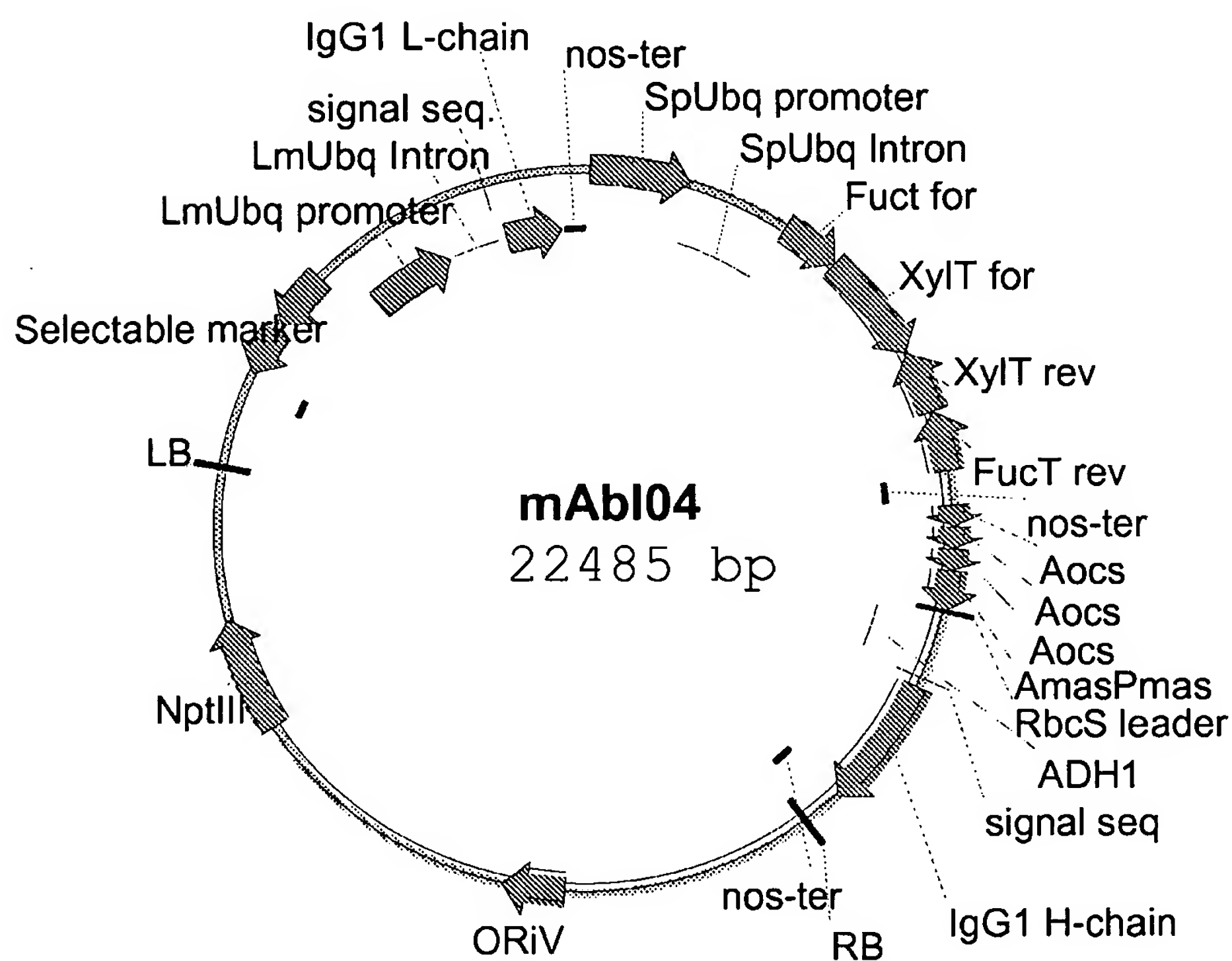
16/59

# **FucT / XylT double knockout Single gene RNAi expression cassette**

**FIG. 11**

17/59

# **IgG1 co-expression with FucT / XylT double knockout (Chimeric RNAi) #1**



**FIG. 12**

18/59

## IgG1 co-expression with FucT / XylT double knockout (single gene RNAi)

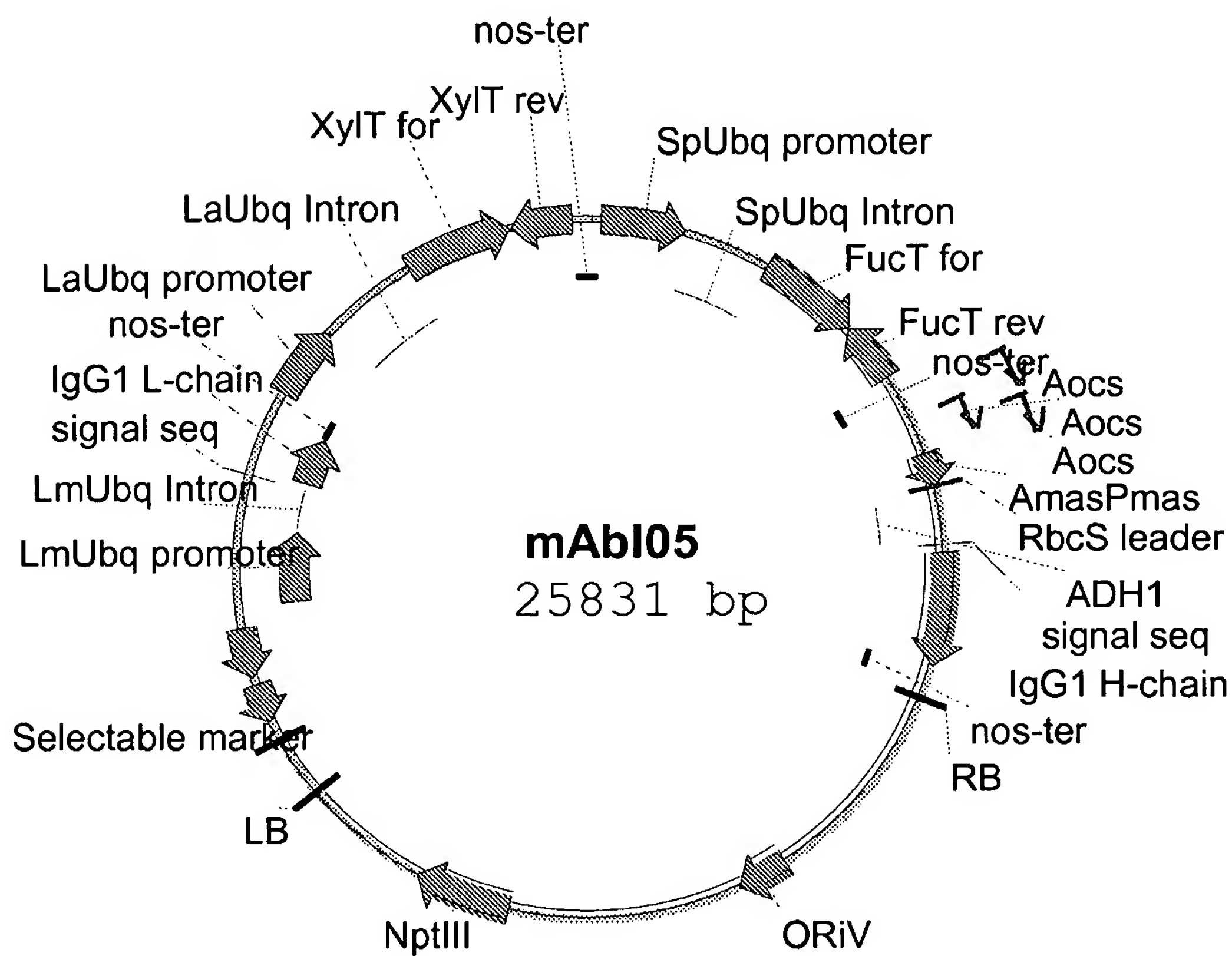


FIG. 13



19/59

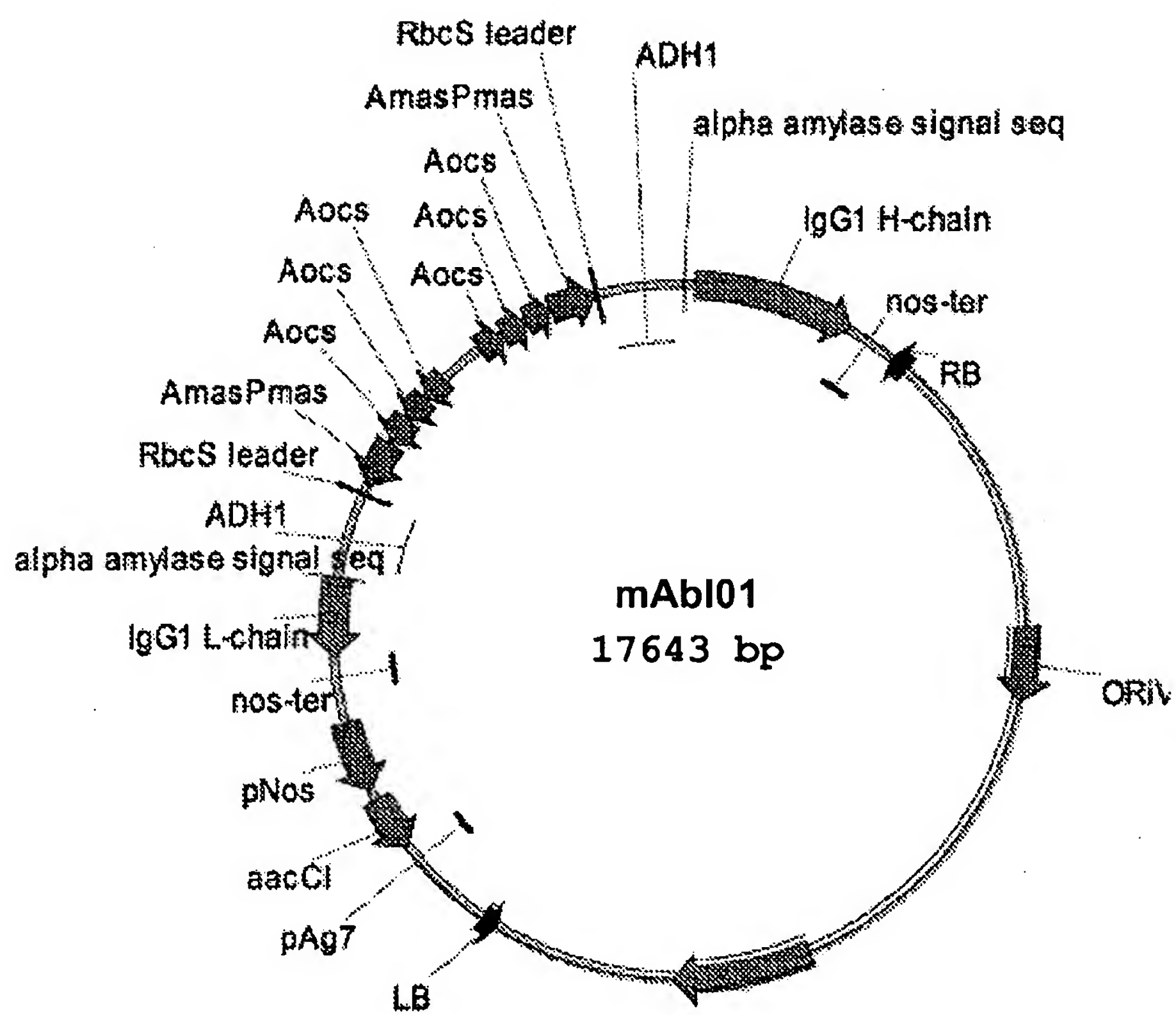
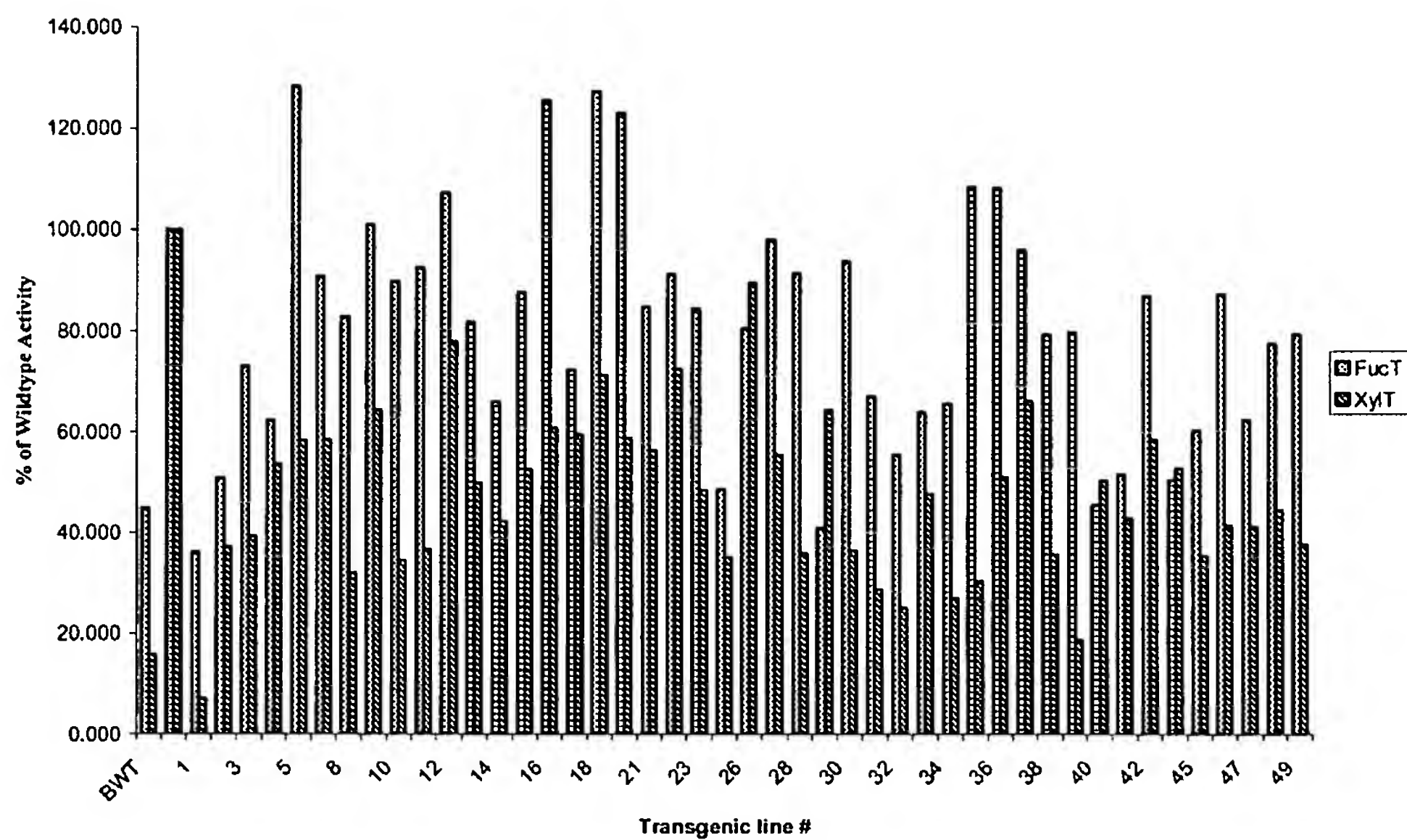


FIG. 14

20/59

## XF02 Set 1 primary screen



## XF02 Set 2 primary screen

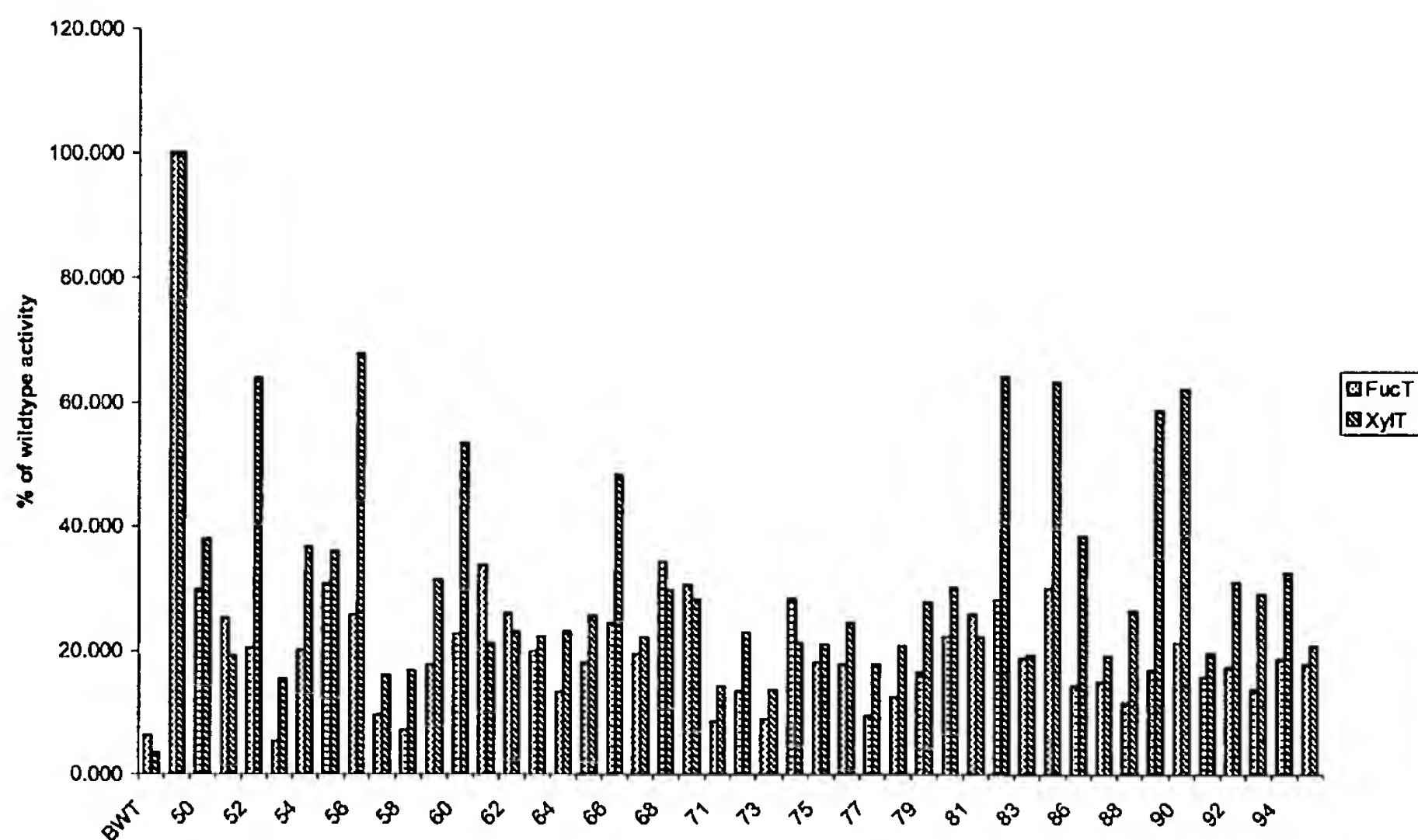


FIG. 15

21/59

## XF02 Set 3 primary screen

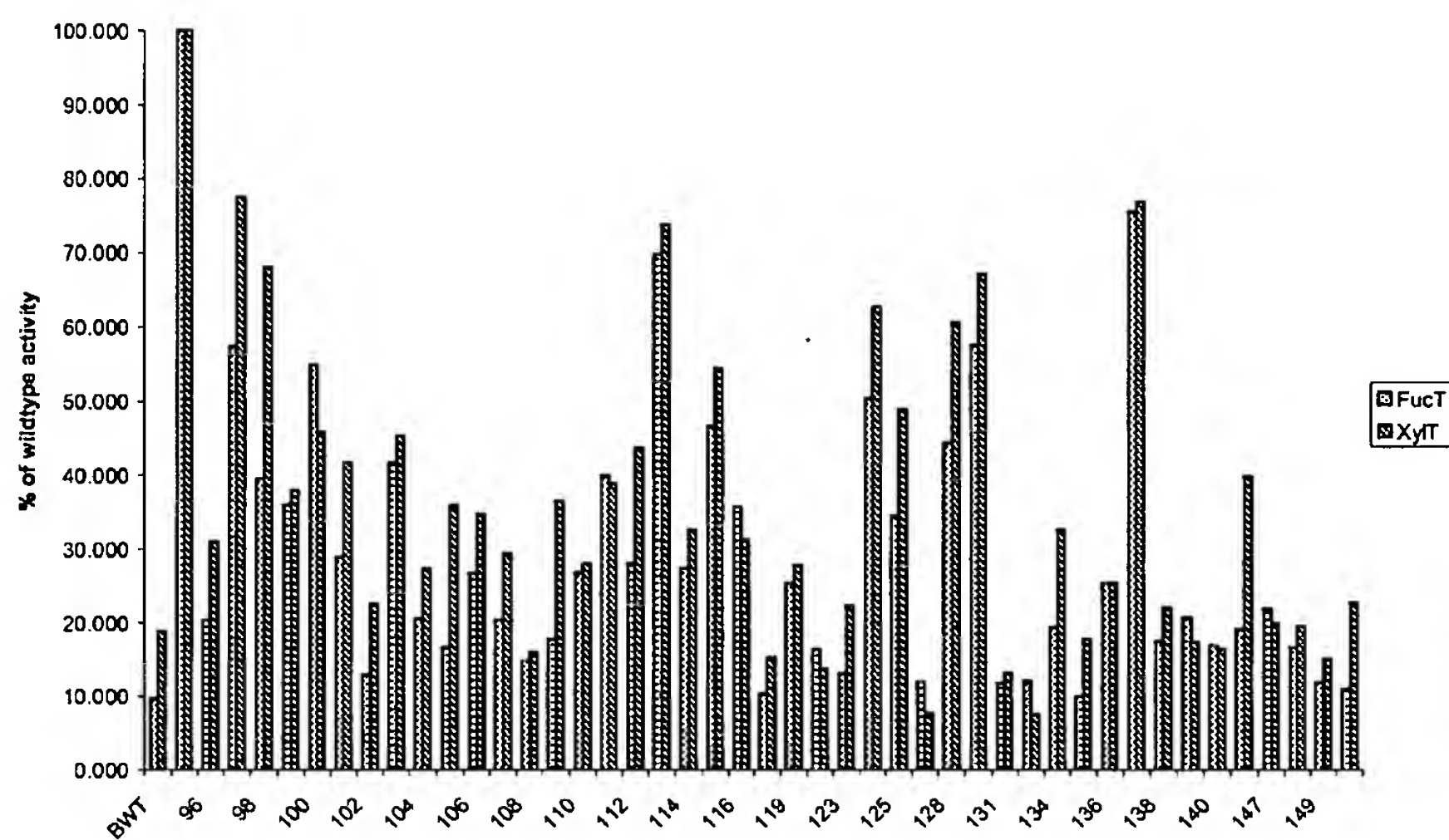


FIG. 16

22/59

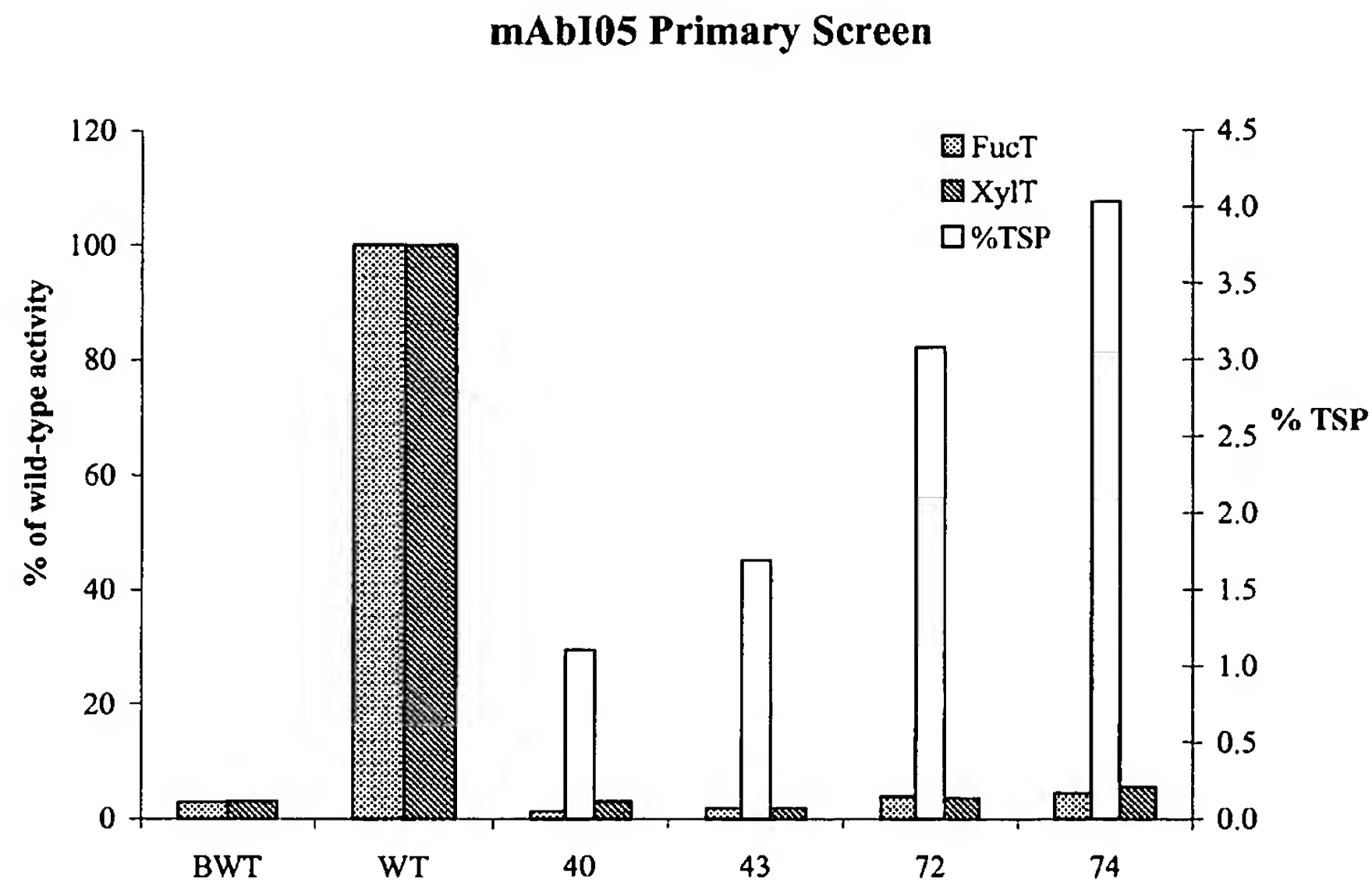
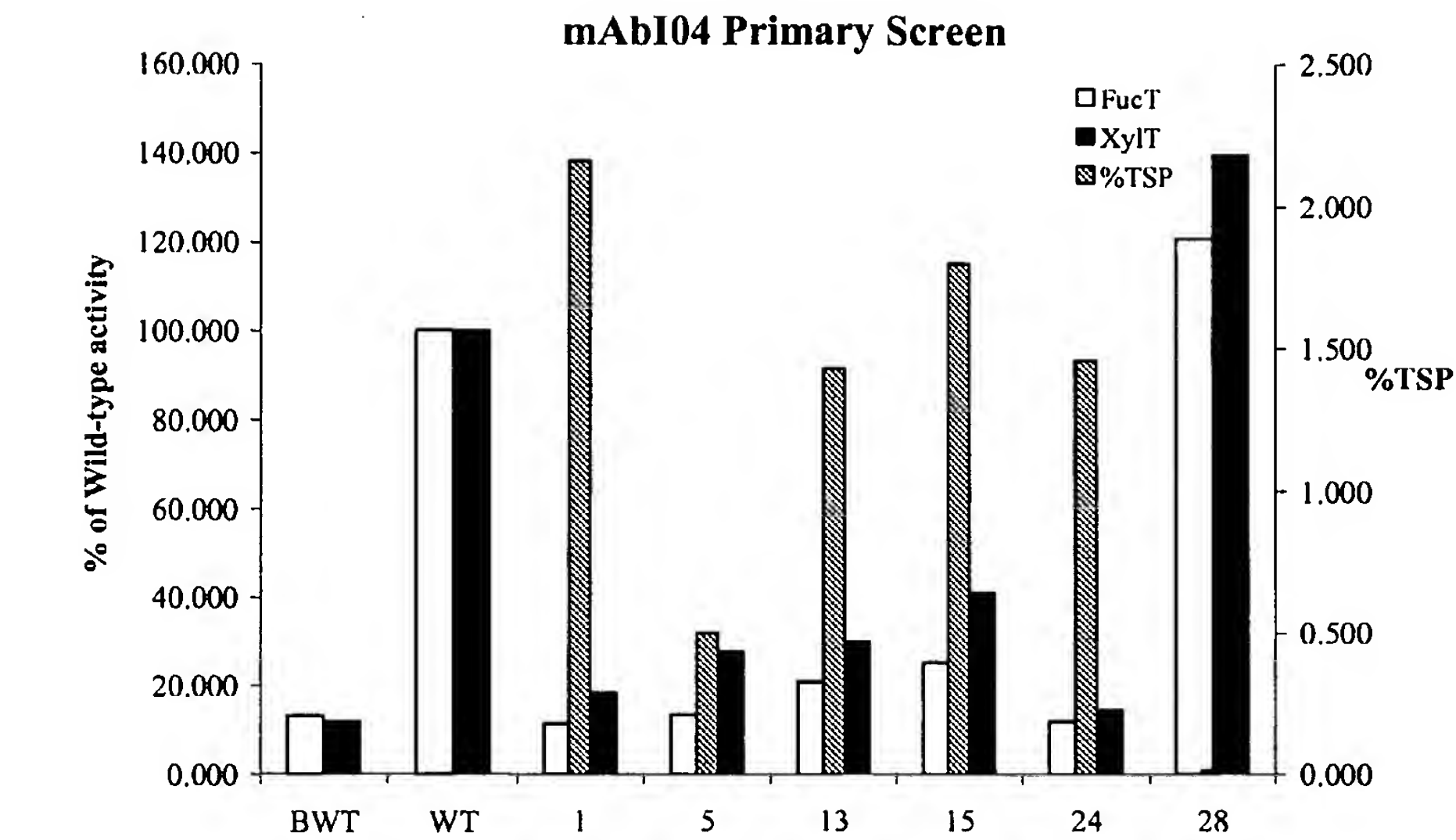


FIG. 17

Structure and molecular weight of derivatized wild-type *L. minor*  
mAb N-glycans.



FIG. 18

Wild-type mAbI01 has three major N-glycans

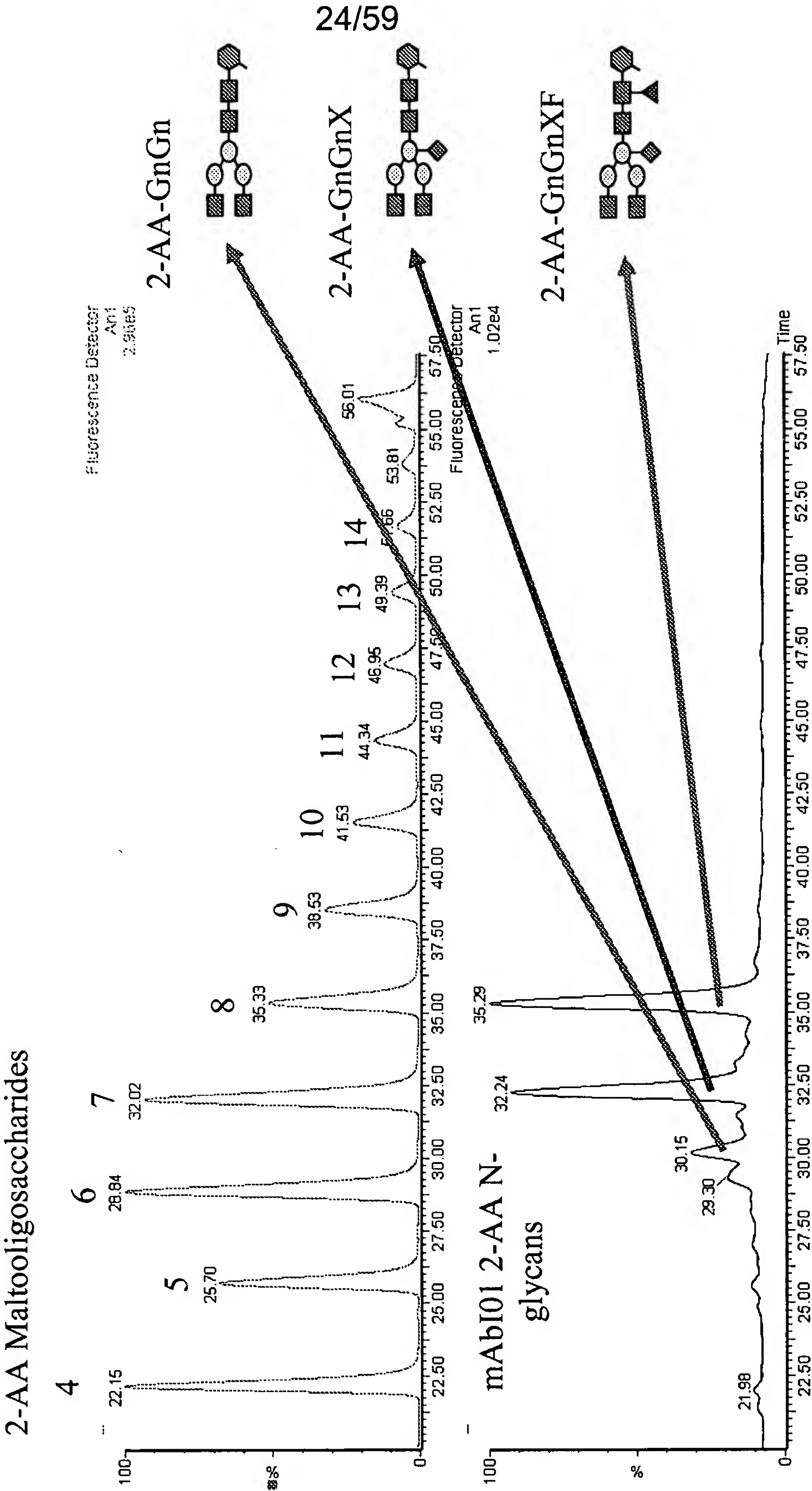


FIG. 19



# LC-MS of 2-AA N-glycans from mAbI01

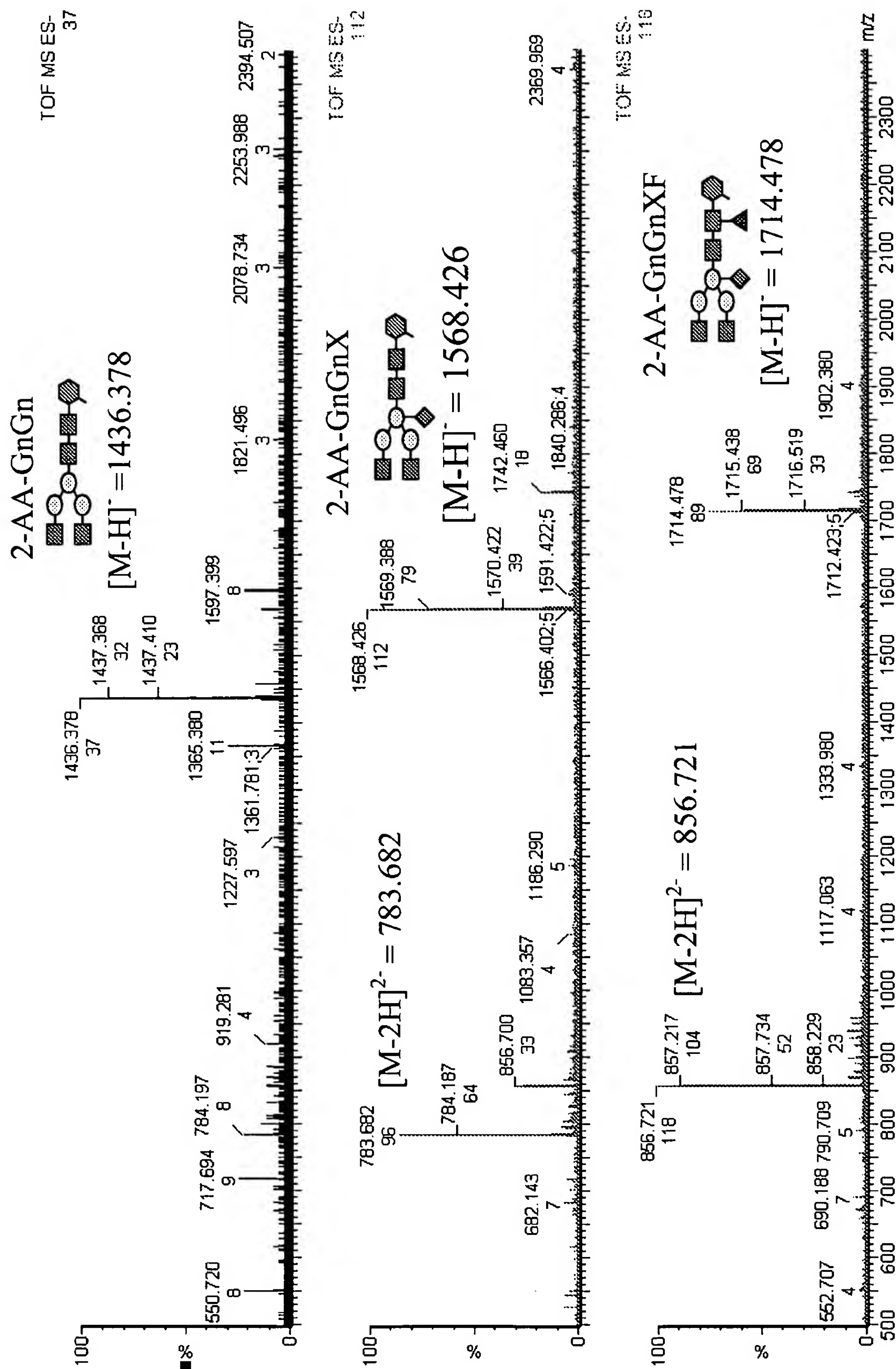


FIG. 20

# MALDI analysis of mAbI01(WT) N-glycans confirms the LC-MS results

mAbI01

26/59

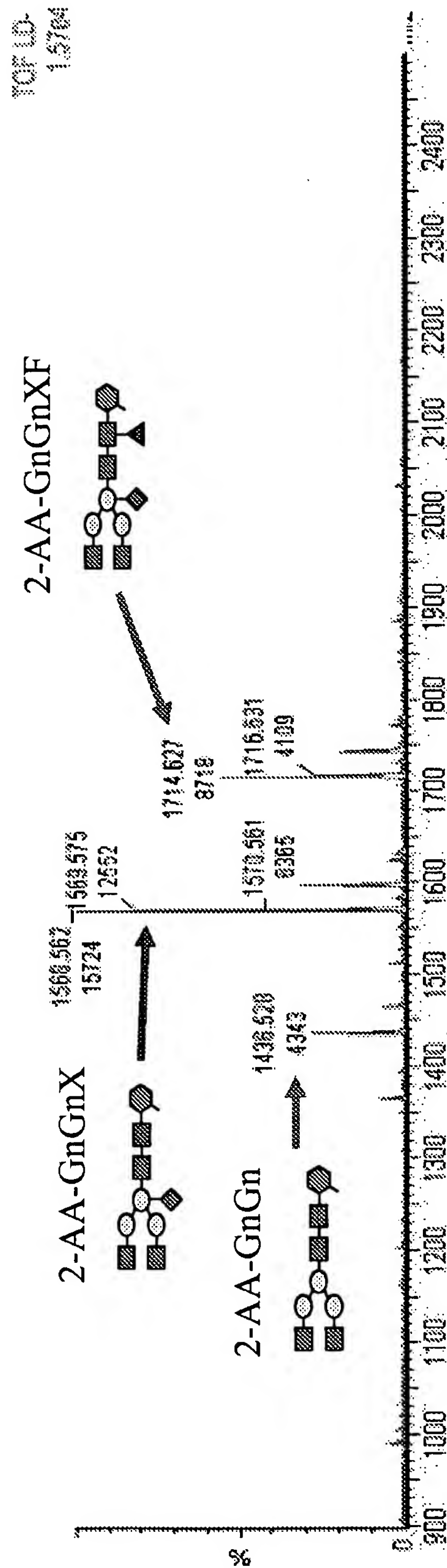


FIG. 21

# Overlay of mAbI01 and mAbI04 N-glycans

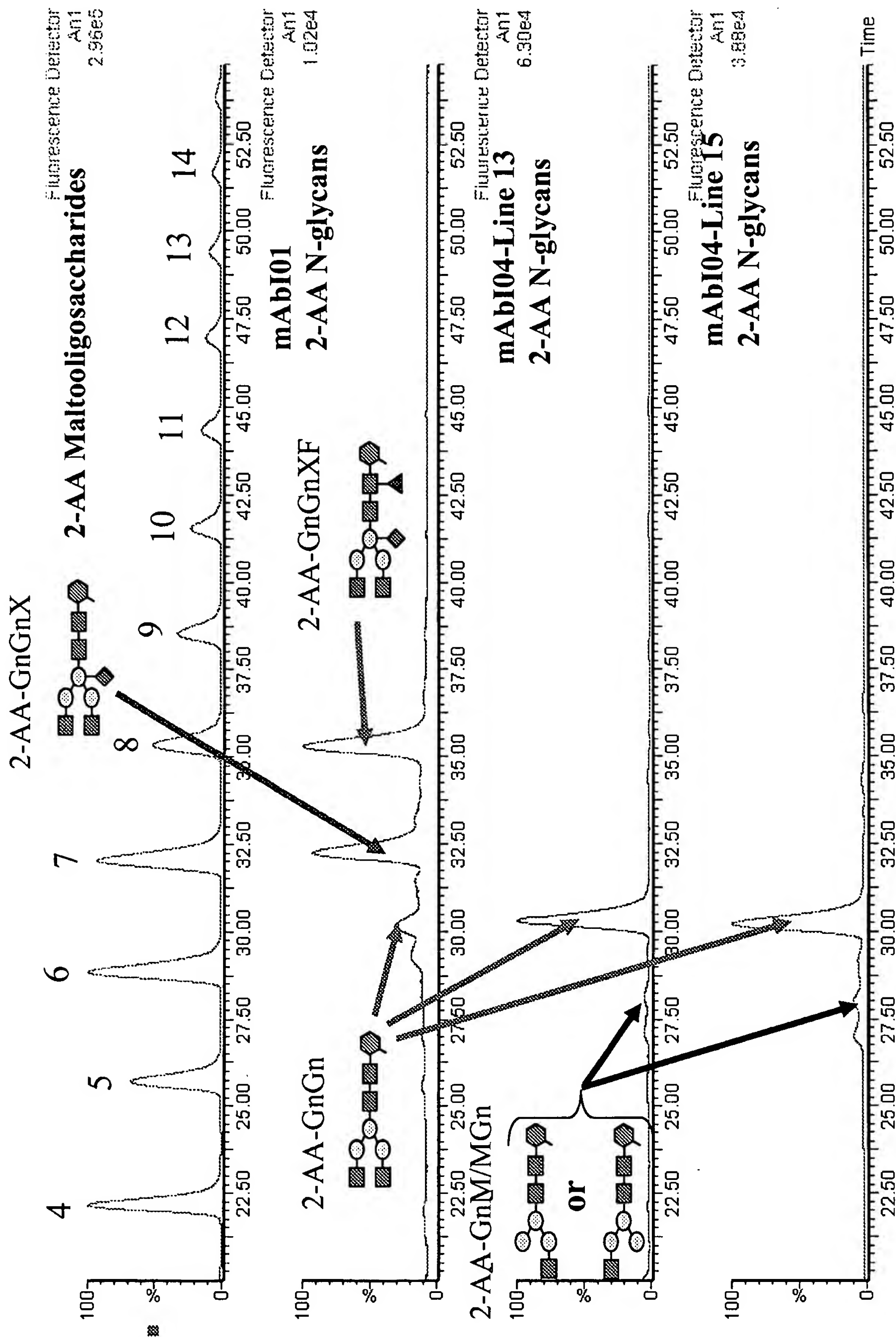


FIG. 22

# LC-MS of 2-AA N-glycans from mAbI04 lines

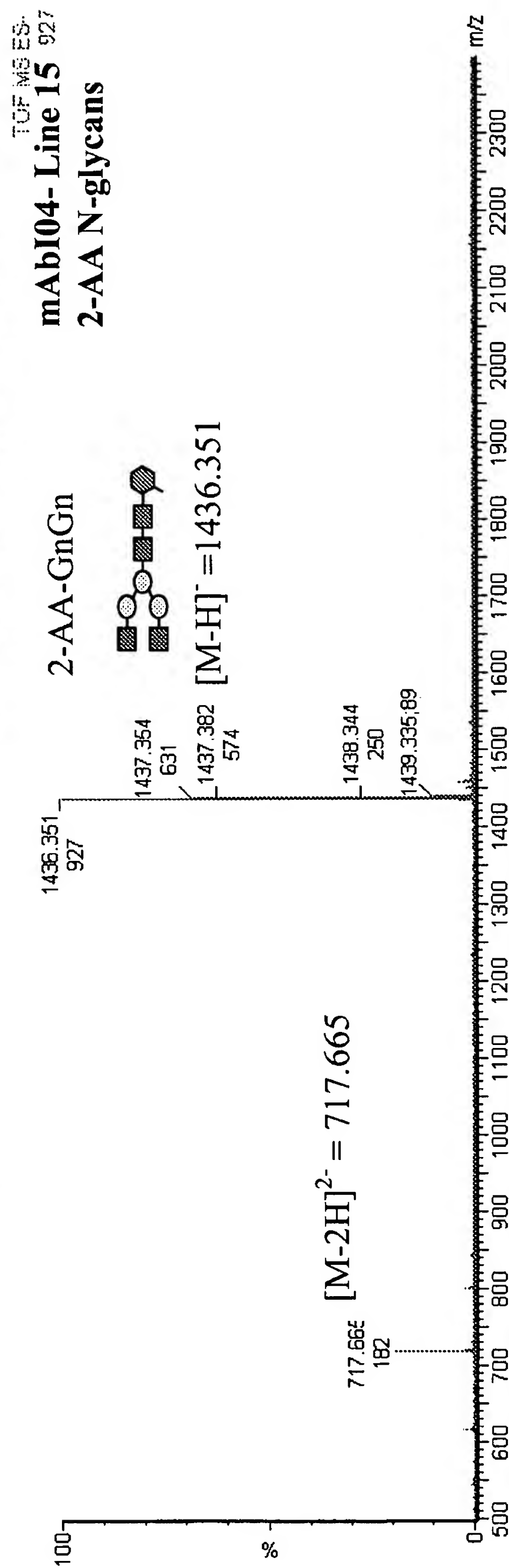
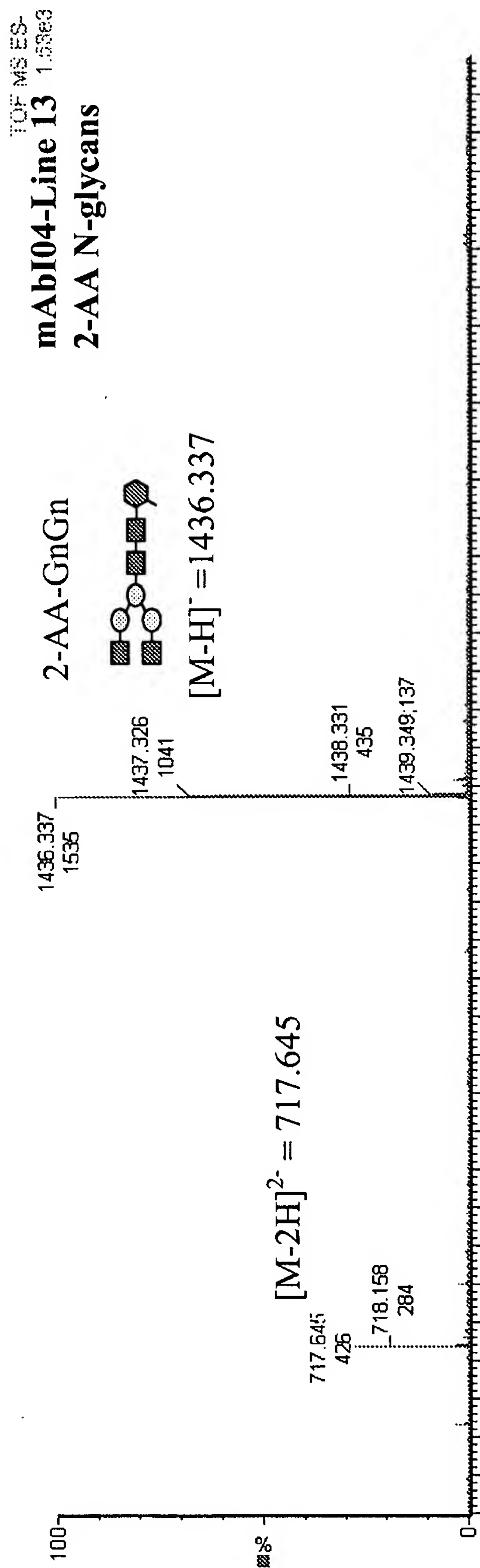


FIG. 23

# MALDI analysis of mAbI01 and mAbI04 N-glycans

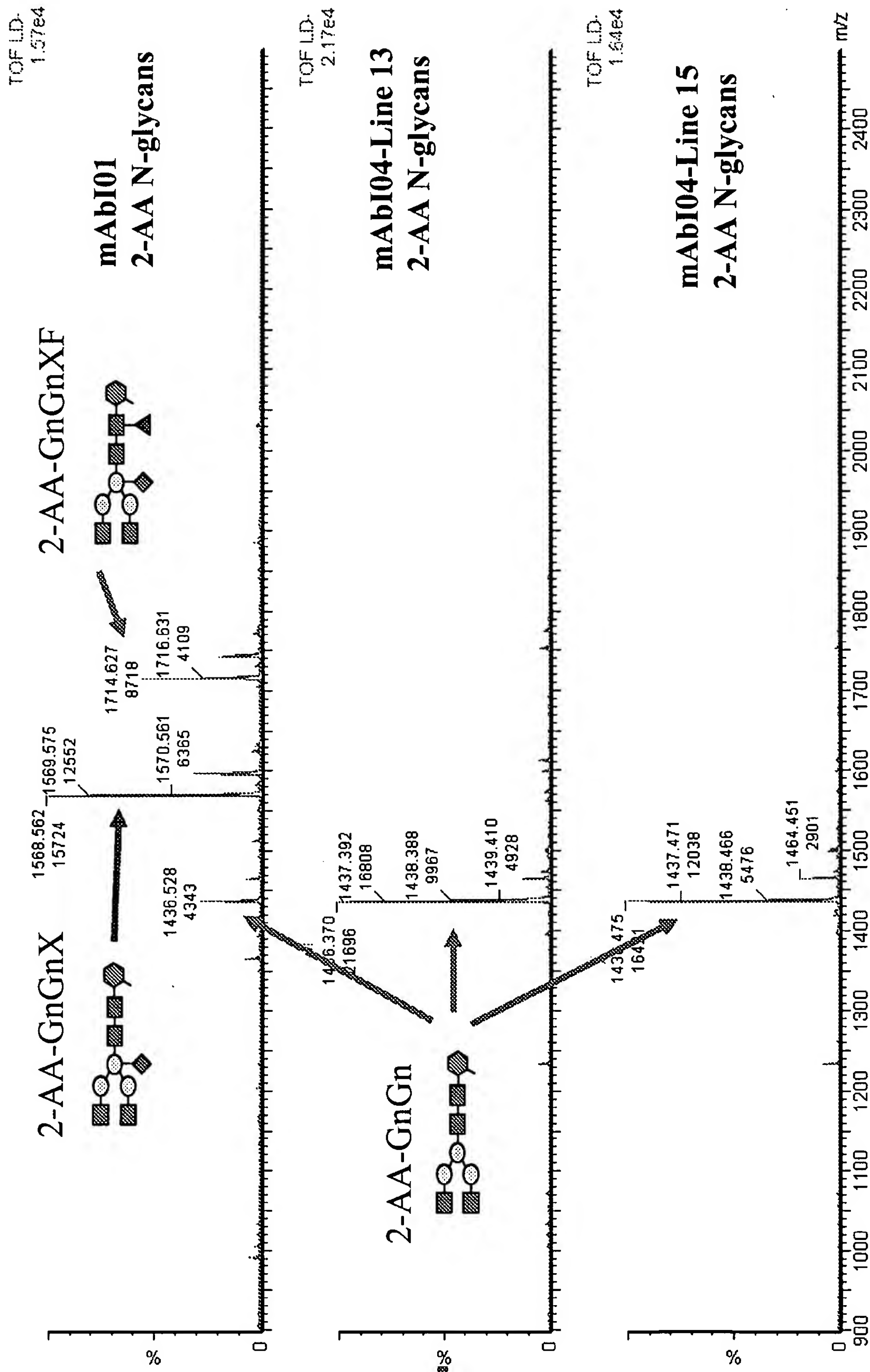


FIG. 24

# Intact mass analysis of mAbI01-20

WO 2007/084926

PCT/US2007/060646

30/59

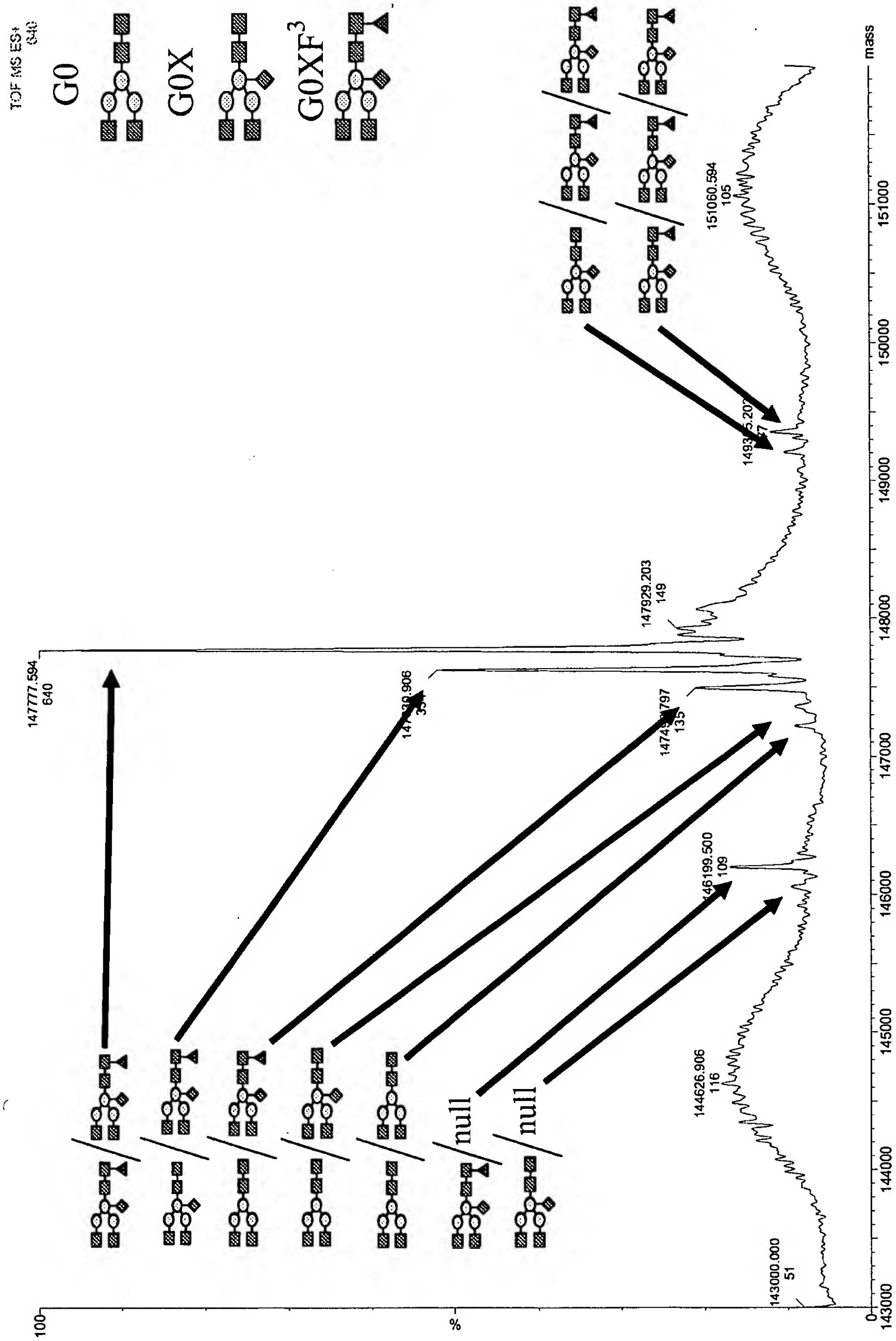


FIG. 25



# Intact mass analysis of mAbI04-15

WO 2007/084926

PCT/US2007/060646

31/59

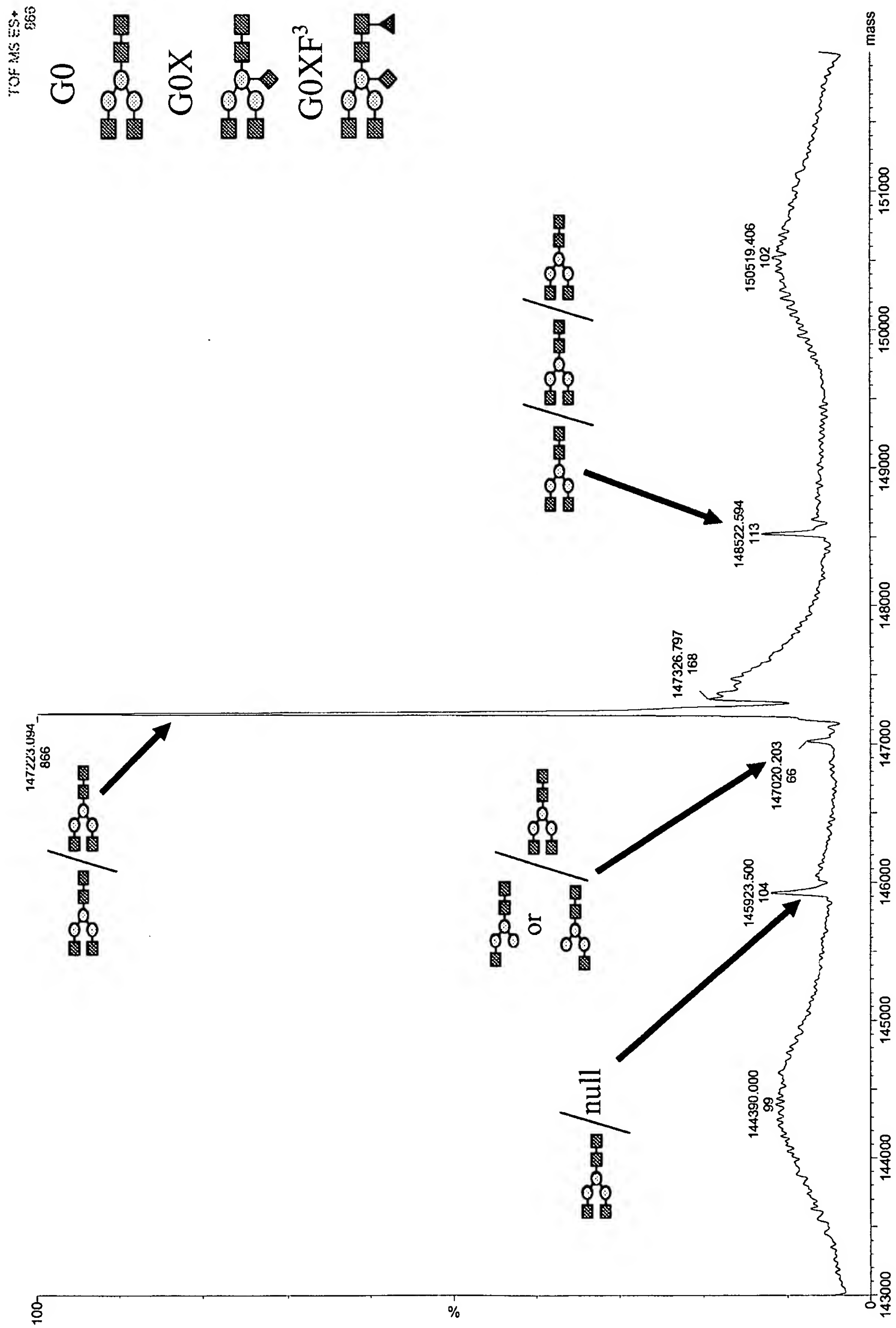
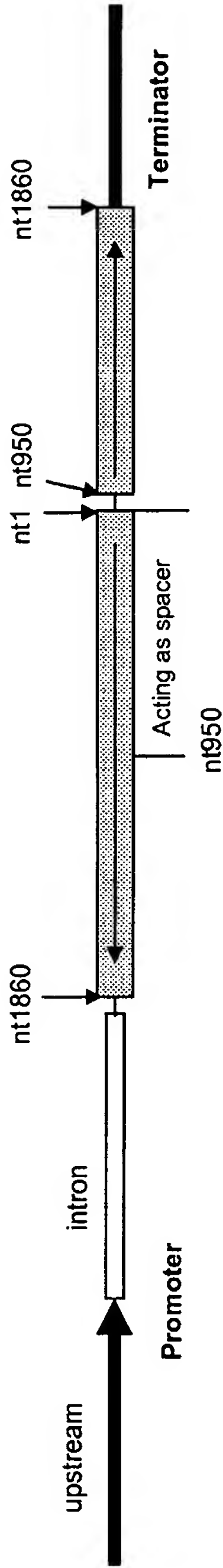


FIG. 26



**Possible design 1 (For targeting individual genes; applicable for both sequences):** first clone the full-length sequence in the antisense orientation, then clone the 3' half of the sequence downstream of the full-length sequence but in the sense orientation. In this way, the 3' half of the sequence (e.g. nt 950-nt1860) will form the dsRNA stem in the hpRNA, and the 5' half of the sequence will act as a spacer. The 3' region is chosen as this region is relatively conserved among different plant species (especially for FucT) compared with the 5' region.



**Possible design 2 (for targeting both genes):** Fuse a ~500 bp sequence from FucT with a ~500 bp sequence from XylT, and then use the fusion sequence to prepare the hpRNA constructs shown below. The sequence between nt700 to nt 1400 represents a region (especially for FucT) that is relatively conserved among different species of plants and is thus representative of one potentially good target. Both of the two versions shown below can be prepared for use together to potentially maximize silencing as preliminary results show that the sequence adjacent to the spacer appears to give better silencing than the distal sequence in the hpRNA construct.

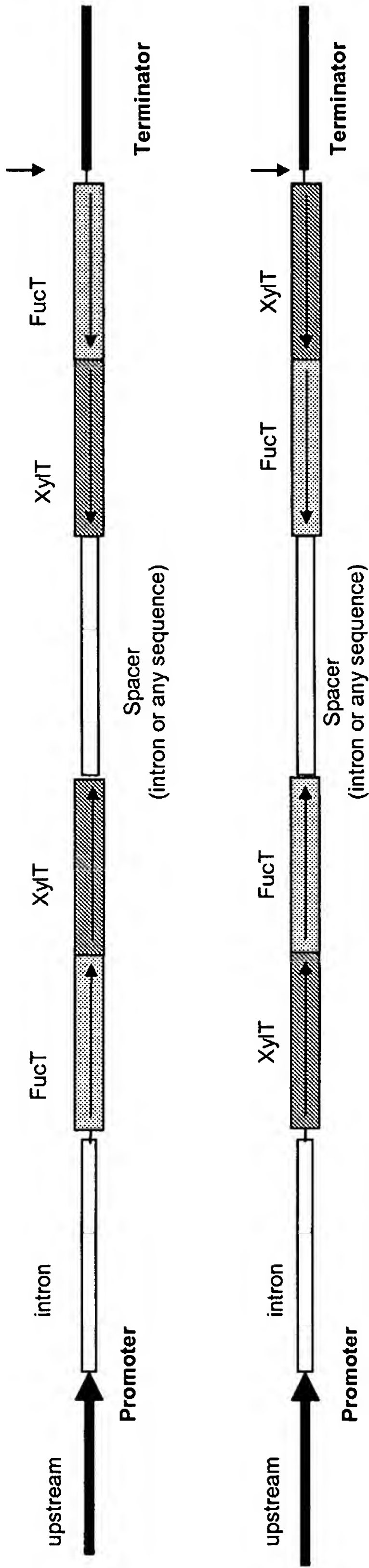


FIG. 28

34/59

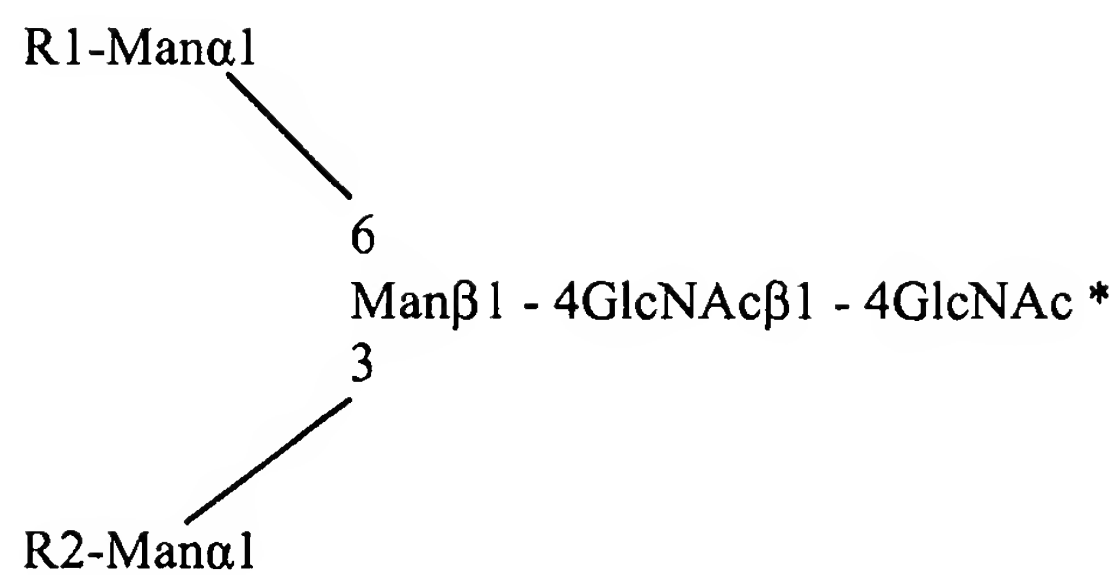


FIG. 29A

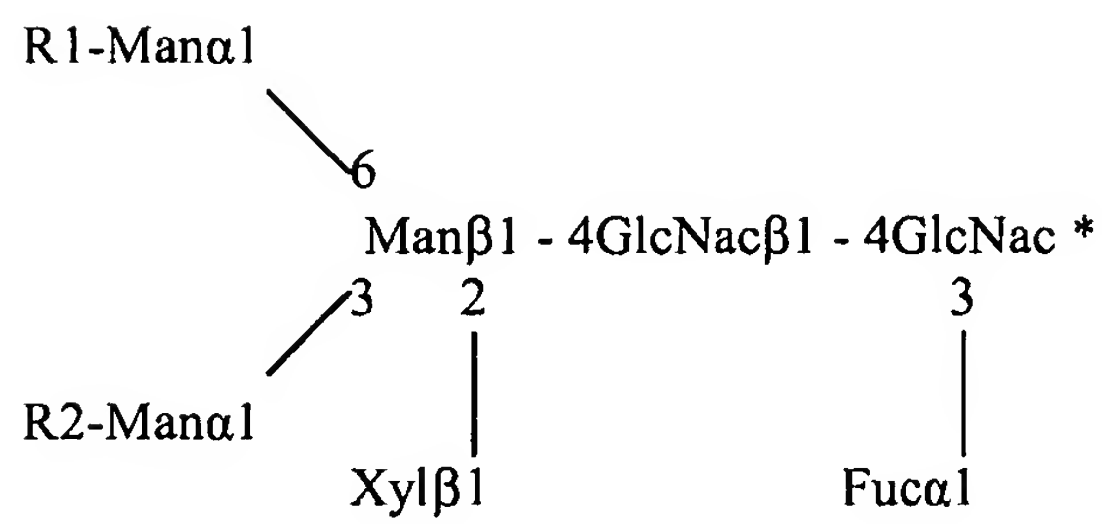


FIG. 29B

35/59

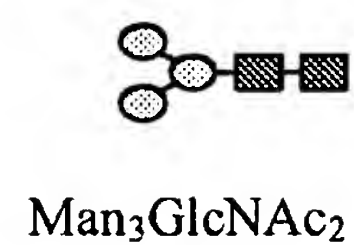
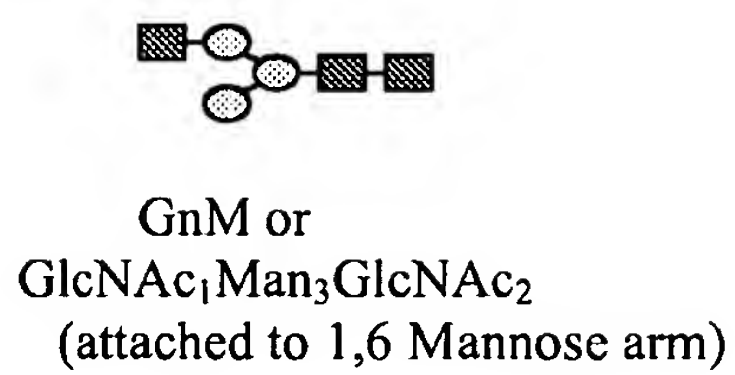
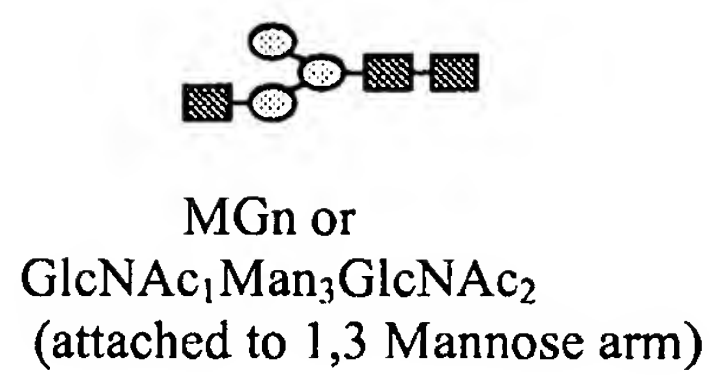
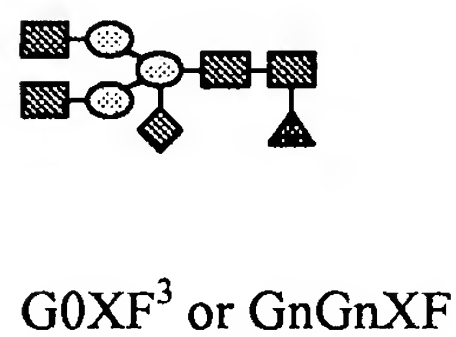
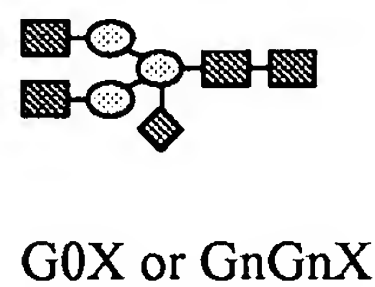
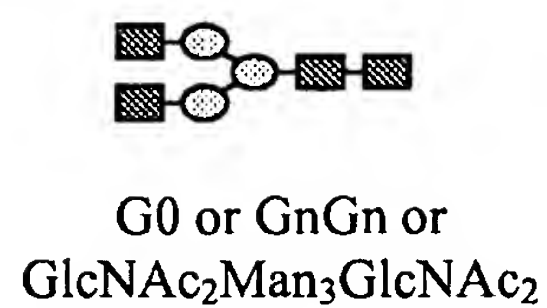
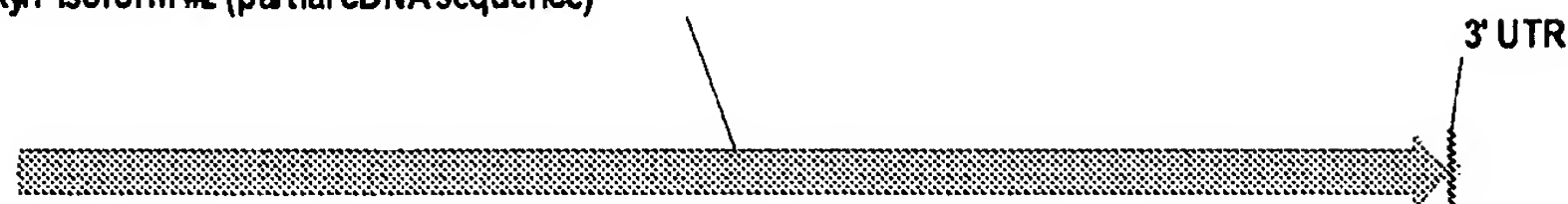
N-GLYCAN STRUCTURELEGEND

FIG. 30

36/59

*Lemna minor* beta 1-2 xylosyltransferase (isoform 2)

XylT Isoform #2 (partial cDNA sequence)



XylT (isoform 2) used in RNAi constructs

1282 bp

XylT isoform #2 partial cDNA sequence (coding sequence in bold font, 3'UTR sequence in standard font):

tgccaagcctacttggccaacagcttcaaccgccggacggagatgctgaagaaggtagaggggaagaggatggttccagtgcctgtaca  
gcgatactcttcgaagttctgtttgccagggagggaatttgcggatggacccggaaaggattaggatgtcgaaagggggggaagatcta  
gaggaggtgatgaagagagaggaggaagaagaattgcccaaatcgaggaggggtcgttccagattgaatctggttatggaagcgga  
ggggaagttggagagagaattgcgactgacgaggtctcgataatgttgtgccgaaggcgctgttcattgtacataccatgcgcaatctc  
atcagttcgattcagattgttgggtccgggcatcttcaatgctctcagtggtcgacgaaccgggttcttctgtcacacgcttcgaatacgcc  
aatctctttcacaccgtcaccgactggtacagcgccctacgcaagctcgaggattgccaaacttgcctctcgccctcacttgattttcgtcgat  
ggccattgcagggcagaacagtttagaggacacgtggcgagccctgttctcaaccgtccgatacgccaagaacttctcccagccagttctgc  
ttccgccacgcegtctctctccctcttggctatgagacagctctcttcaaaggcctatcagagagcttcagctgtgagggagtgccggccaa  
tcagctcaaagtcaaccctgatgaccagaagactgagagactggctgaattcggagagatgatcagggctgcctttgactttctgtcgtt  
gacccgcccgttgacccgttgaccaaattccatctcttctgtgcggcgagggaagattacgtggcgccaccacgccacagtgaggagagtgag  
tcgggttgaccaatgagcaagagggtgtttgactttctgcacaaatgggcaagtcaacacagaagcaggtgcaacgtcagtggtgcaac  
gggcttttcgggcacatgggaatgaaggaaacagggtgaaggcaattatggaagcttcgggtggtggtcgggggccacggggciggtttgact  
catctgggtggcagcaagggtcaacgacagttgttcttgagattctgagcagtcataatcgtagaccgcactttcaactgatttcacgggtgga  
agggttggactaccacgcaattaatcttgcggggtcgatgctgatctctcgagggtggtcgagaaattgactggcatagtcgatgggctt  
ggatgttgaagataag

XylT isoform #2 predicted amino acid sequence (partial sequence):

ceayfgnsfnrrtemlkkvegrgwfqclysdtrssvcqggnlrm dperirmskggedleevmkreecelpkfeegsfqiesgygsggev  
geriatdevldnvvpkgavhvhtmrnlissiqivgpghlqcsqwdepvllvtrfeyanlfhtvtdwysayassrianlpsrphlifvdghcrae  
qledtwralfstvryaknfsqpvcfrahvlsplgyetalfkglsefscegvpanqlkvnppdqktarlaefgemiraafdfpvvdppvdpltk  
ilfvrredyvahprhsgrvesrltneqevfdflhkwasqhrsncvsvvnglfahmgmkeqvkaimeasvvvgahgagtlhlvaarsttvl  
eilssqyrphqlisrwkgl dyhainlagsyadprevvekltgivdglgc\*

FIG. 31

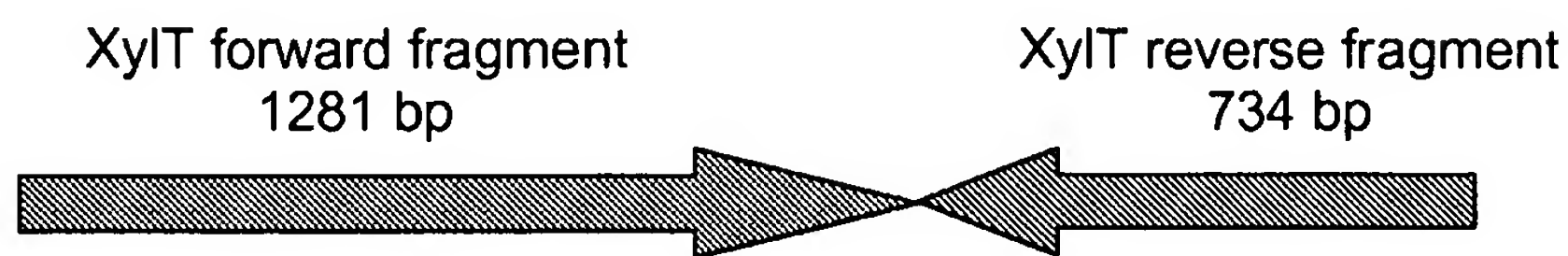


37/59

		1	40
LmXylT Isoform#1	(1)	MALVNSRGSRRVRIAKPTFVFLINVVCLLYFFRQNPNI	
LmXylT Isoform #2	(1)	-----	
		41	80
LmXylT Isoform#1	(41)	PDACLHGEC DKPPILVTPRRWNLKPWPILPSFLPWVPSSH	
LmXylT Isoform #2	(1)	-----	
		81	120
LmXylT Isoform#1	(81)	PAQGSCEAYEGNSFNFFTEMLKKVEGRGWFOCLYSOTLPS	
LmXylT Isoform #2	(1)	-----CEAYEGNSFNFFTEMLKKVEGRGWFOCLYSOTLPS	
		121	160
LmXylT Isoform#1	(121)	SVCQGGNLRMDPERLRMSKGGEDLEEVMKREEEELPKKE	
LmXylT Isoform #2	(36)	SVCQGGNLRMDPERLRMSKGGEDLEEVMKREEEELPKKE	
		161	200
LmXylT Isoform#1	(161)	EGSFQIESGYGSGGEVGERIATDEVLDNVVPGAVHVRHM	
LmXylT Isoform #2	(76)	EGSFQIESGYGSGGEVGERIATDEVLDNVVPGAVHVRHM	
		201	240
LmXylT Isoform#1	(201)	RNLISSTQIVGPGHLCSSQWIDEFVLLVTPREYANLEHTV	
LmXylT Isoform #2	(116)	RNLISSTQIVGPGHLCSSQWIDEFVLLVTPREYANLEHTV	
		241	280
LmXylT Isoform#1	(241)	TDWYSAYASSRIANLPSRPHLEFVDGHCRAEQLEDNRKAI	
LmXylT Isoform #2	(156)	TDWYSAYASSRIANLPSRPHLEFVDGHCRAEQLEDNRKAI	
		281	320
LmXylT Isoform#1	(281)	FSTVRYSPNESQPCFPHVLSPLGYETALFKGLSESEFSC	
LmXylT Isoform #2	(196)	FSTVRYSPNESQPCFPHVLSPLGYETALFKGLSESEFSC	
		321	360
LmXylT Isoform#1	(321)	EGAFANRLKVRFDQKTARLAEEGEMFAAFDEPVVDEPS	
LmXylT Isoform #2	(236)	EGVFANQLKVRFDQKTARLAEEGEMFAAFDEPVVDEPS	
		361	400
LmXylT Isoform#1	(361)	DPLTKSILFVRKREDAVHPRHSGRVE SRLTNEQEVDEFLH	
LmXylT Isoform #2	(276)	DPLTKSILFVRKREDAVHPRHSGRVE SRLTNEQEVDEFLH	
		401	440
LmXylT Isoform#1	(401)	NWASHHREGRCNLSVNGLEFAHGMKEQSKATMEASVVVGA	
LmXylT Isoform #2	(316)	KWASQHSRCNLSVNGLEFAHGMKEQSKATMEASVVVGA	
		441	480
LmXylT Isoform#1	(441)	HGAGLTHIVAFSTETVLELSSQYRPPHFOLISPWKGLD	
LmXylT Isoform #2	(356)	HGAGLTHIVAFSTETVLELSSQYRPPHFOLISPWKGLD	
		481	510
LmXylT Isoform#1	(481)	YHAINLAGSSADPREVVEKITGIVRLGC-	
LmXylT Isoform #2	(396)	YHAINLAGSSADPREVVEKITGIVRLGC--	

FIG. 32

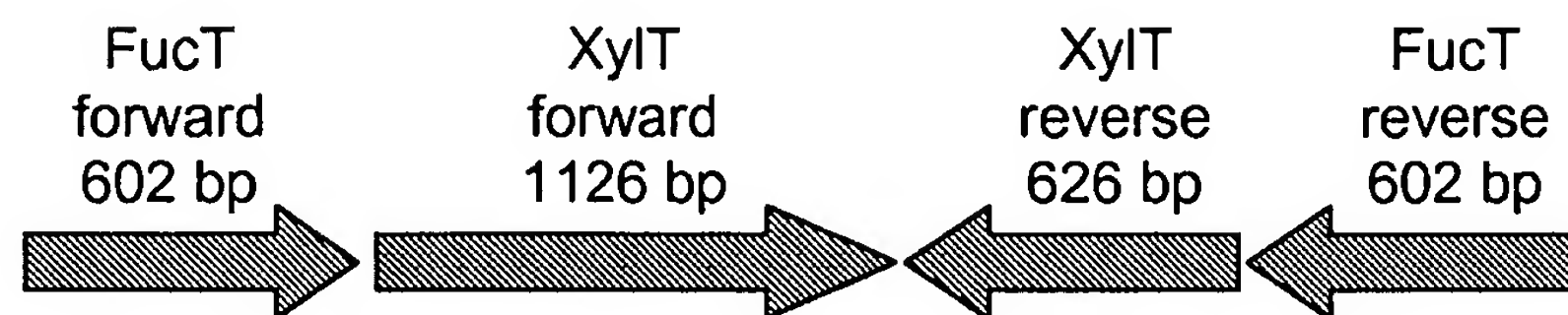
38/59

**XylT RNAi design (single gene KO):**

- Xylosyltransferase (XylT) RNAi expression cassette was designed to express a hairpin RNA with 734 bp of double-stranded RNA sequence and a 547 bp spacer sequence, based on the XylT isoform #2 partial DNA sequence. The RNAi expression cassette comprises the XylT forward fragment (nt 1-734 of SEQ ID NO:19), the spacer sequence (nt 735-1282 of SEQ ID NO:19), and the XylT reverse fragment (antisense version of nt 1-734 of SEQ ID NO:19).

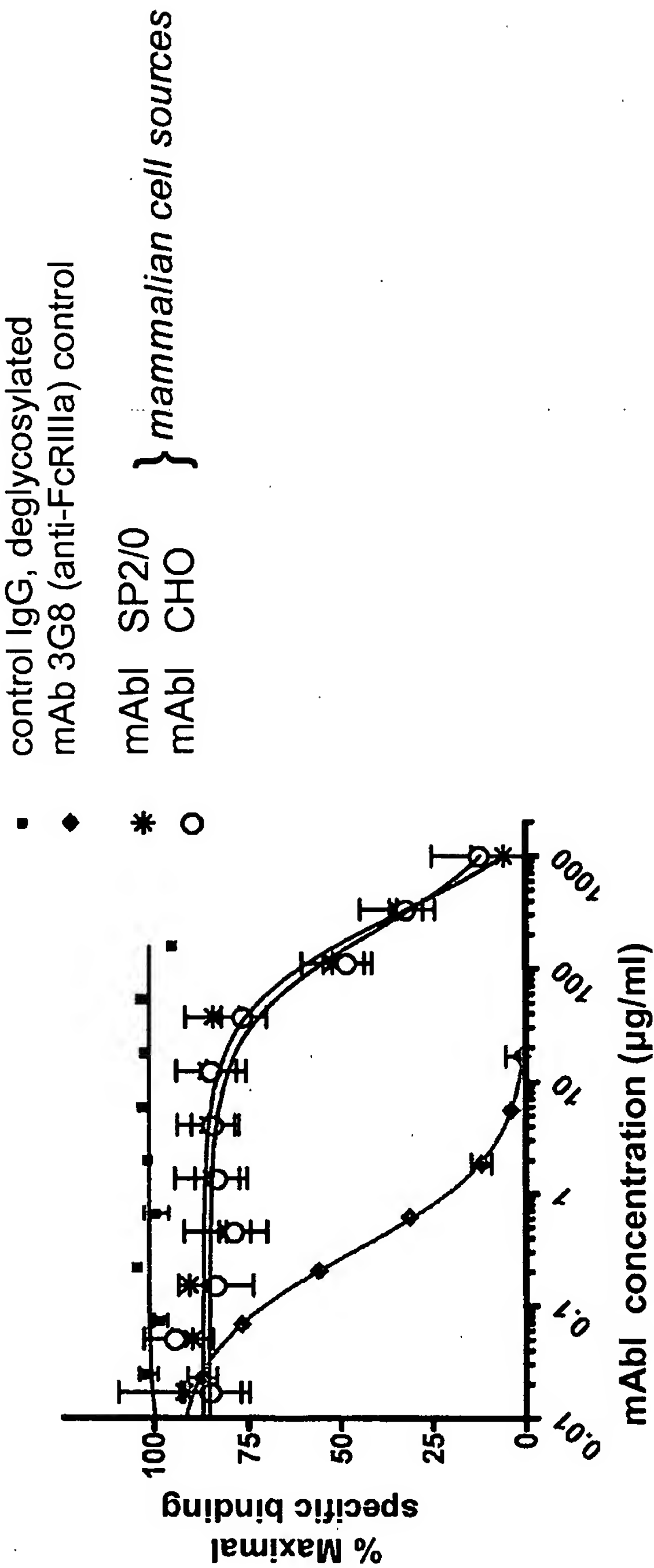
**FIG. 33**

39/59

**FucT + XylT chimeric RNAi design:**

- Chimeric FucT + XylT RNAi expression cassette was designed to express a hairpin RNA with 602 bp of double-stranded FucT sequence, 626 bp of double-stranded XylT sequence, and a 500 bp spacer sequence. The RNAi cassette encompasses in the 5' to 3' direction and operably linked: the FucT forward fragment (nt 254-855 of SEQ ID NO:1), the XylT forward fragment from isoform #2 (nt 1-626 of SEQ ID NO:19), a spacer sequence (nt 627-1126 of SEQ ID NO:19), the XylT reverse fragment (antisense version of nt 1-626 of SEQ ID NO:19), and the FucT reverse fragment (antisense version of nt 254-855 of SEQ ID NO:1).

**FIG. 34**



<sup>125</sup> I-mAb 3G8 (anti-FcγRIIIa) + mAb test Ab
Human NK cells

FIG. 35

41/59

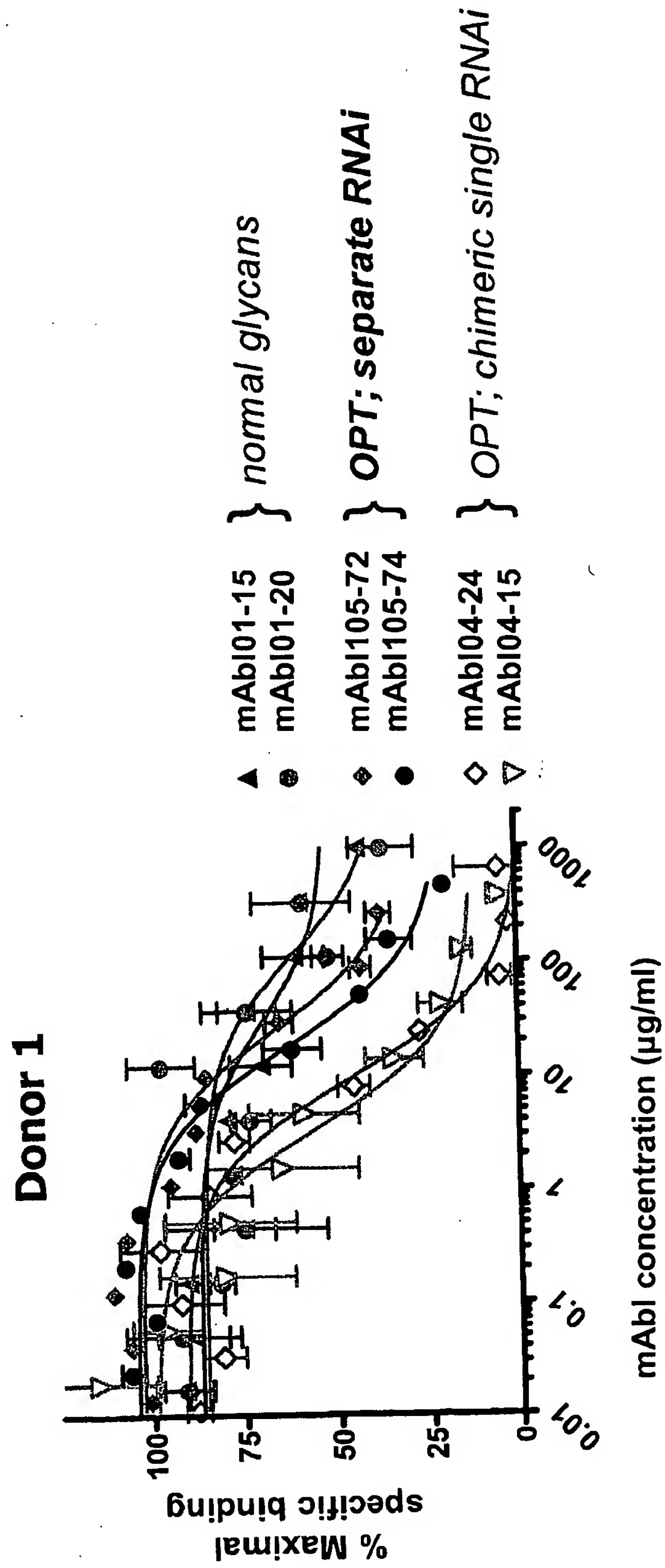


FIG. 36

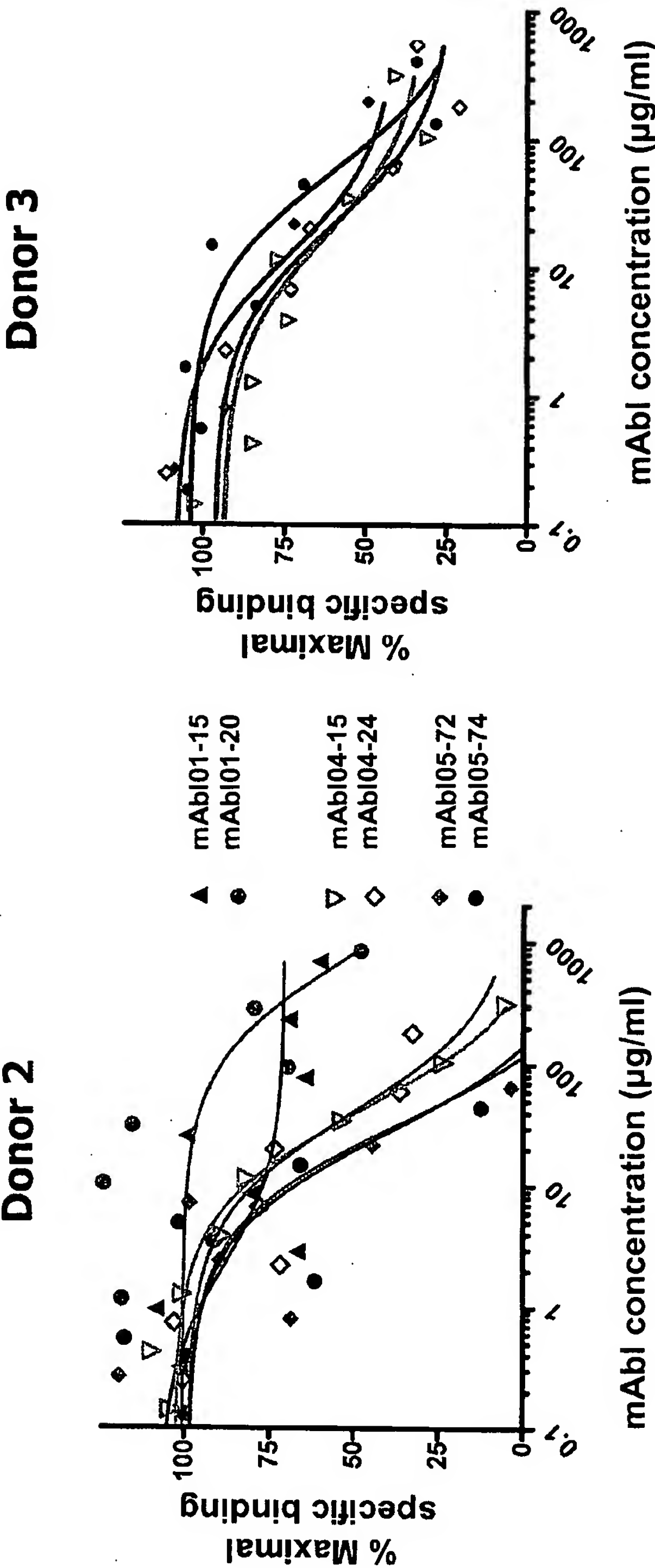


FIG. 37



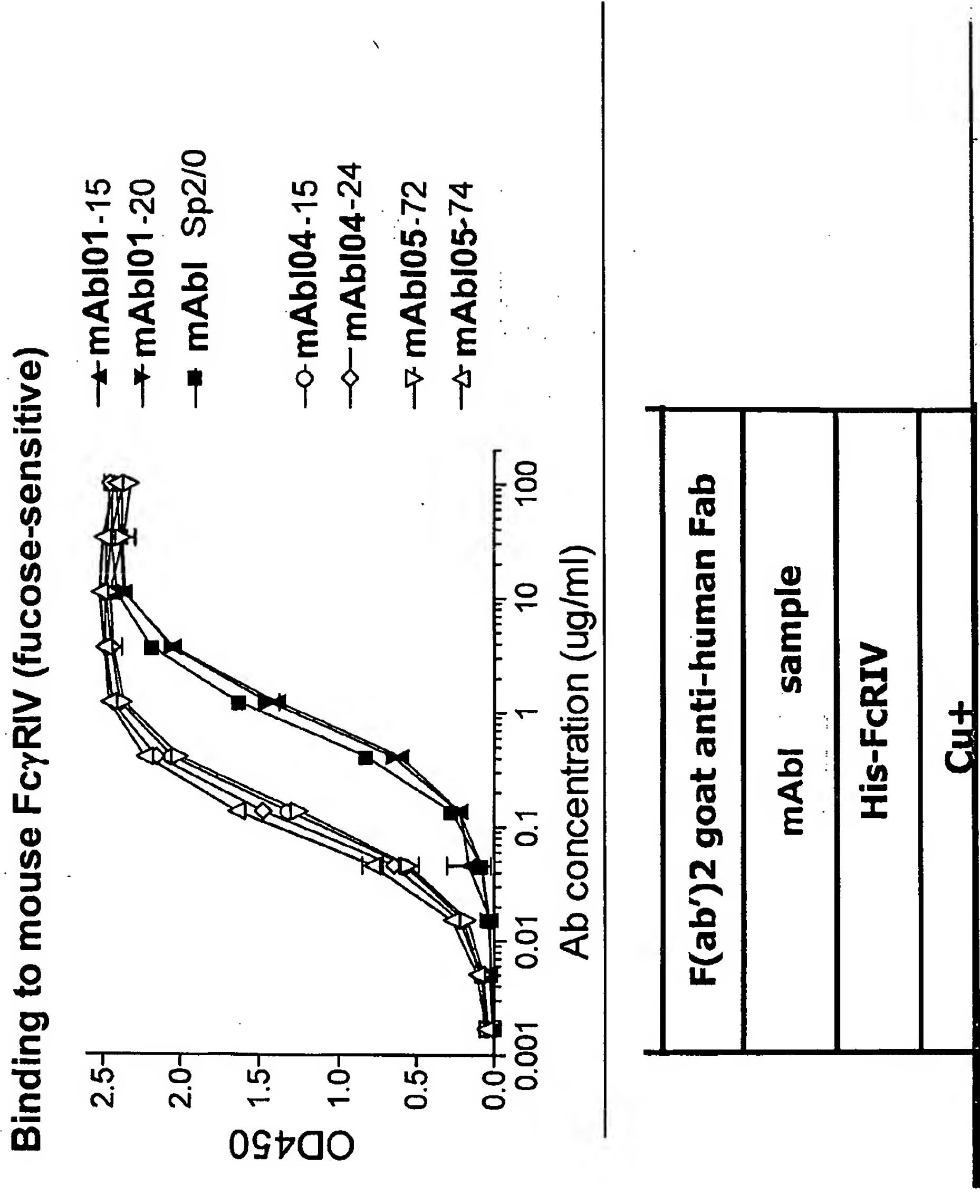


FIG. 38

44/59

## MDXA04 Expression Vector

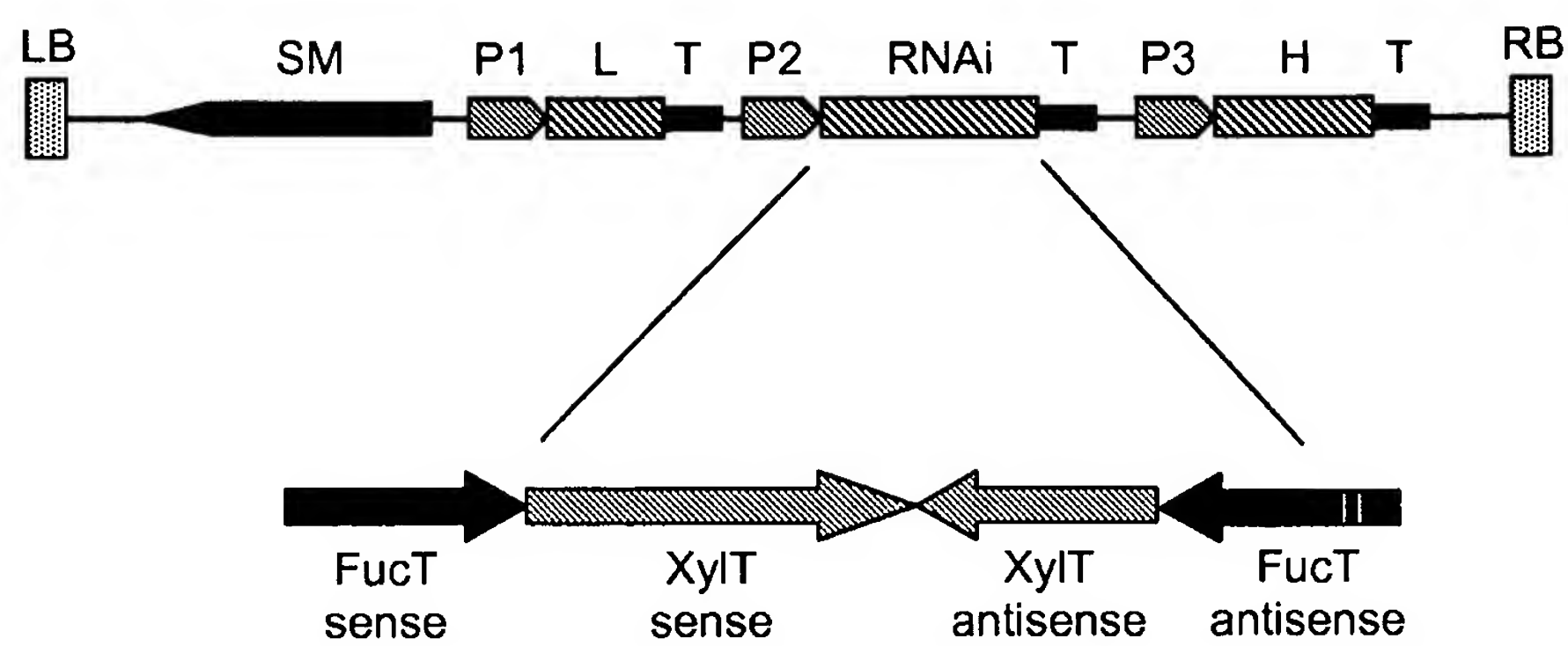


FIG. 39

45/59

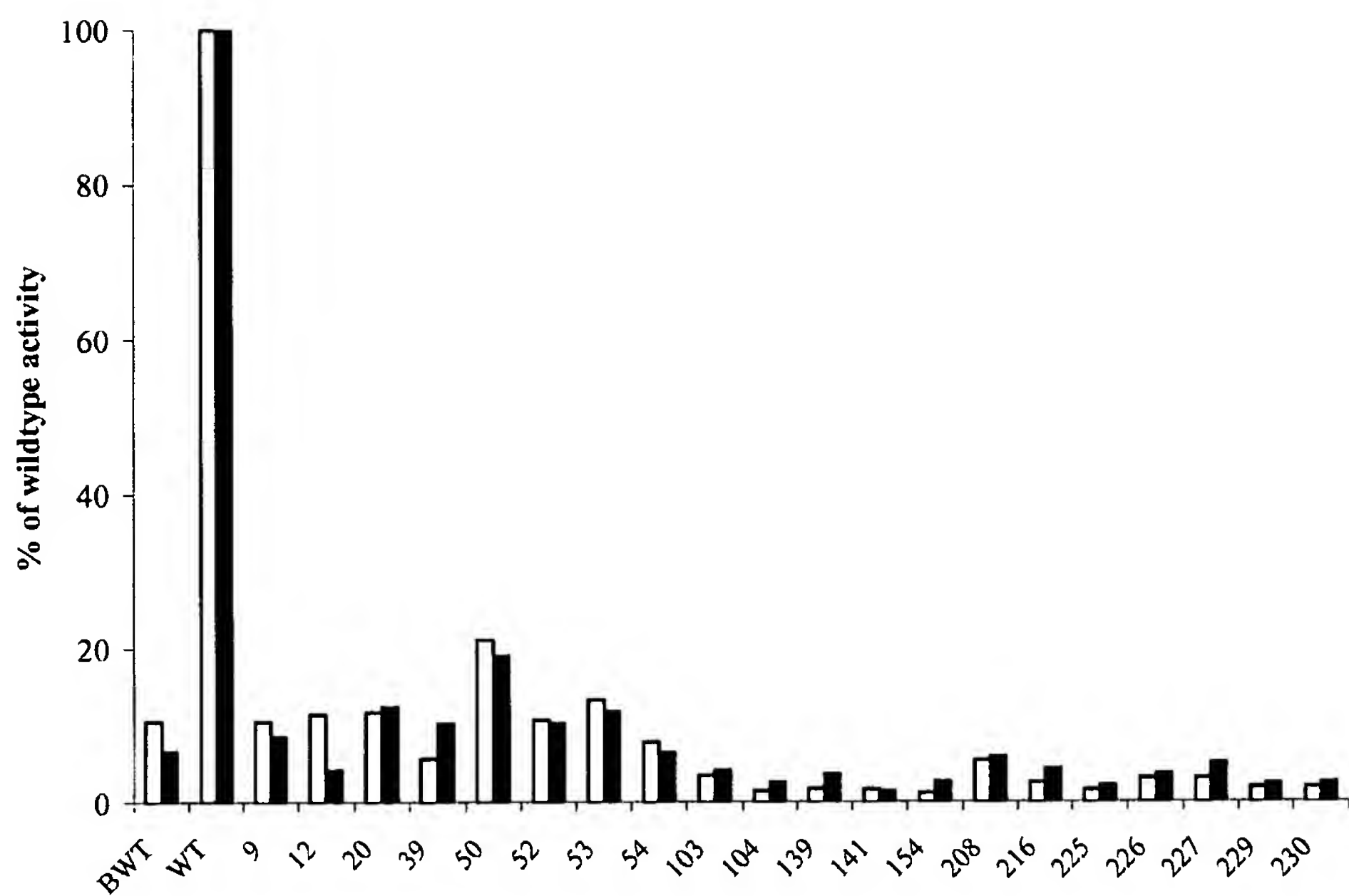


FIG. 40

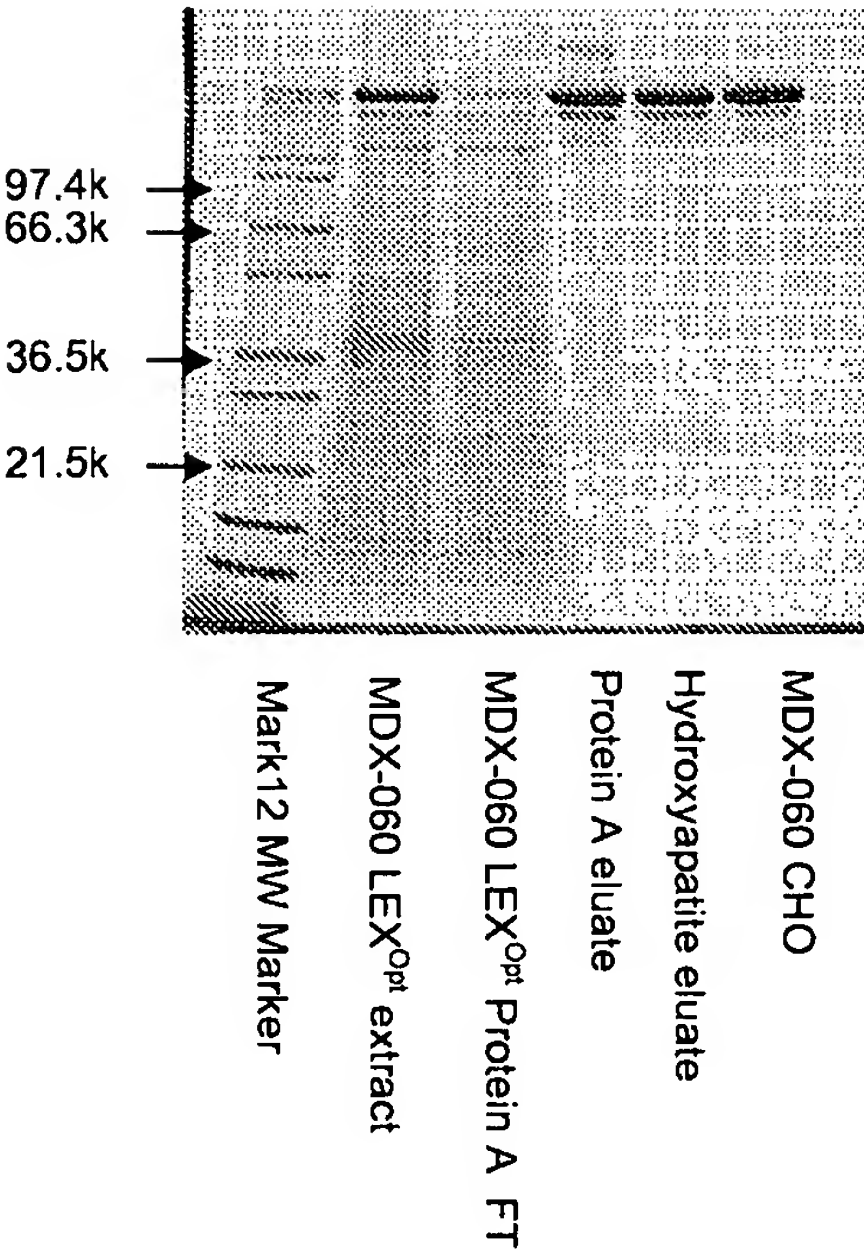


FIG. 41A

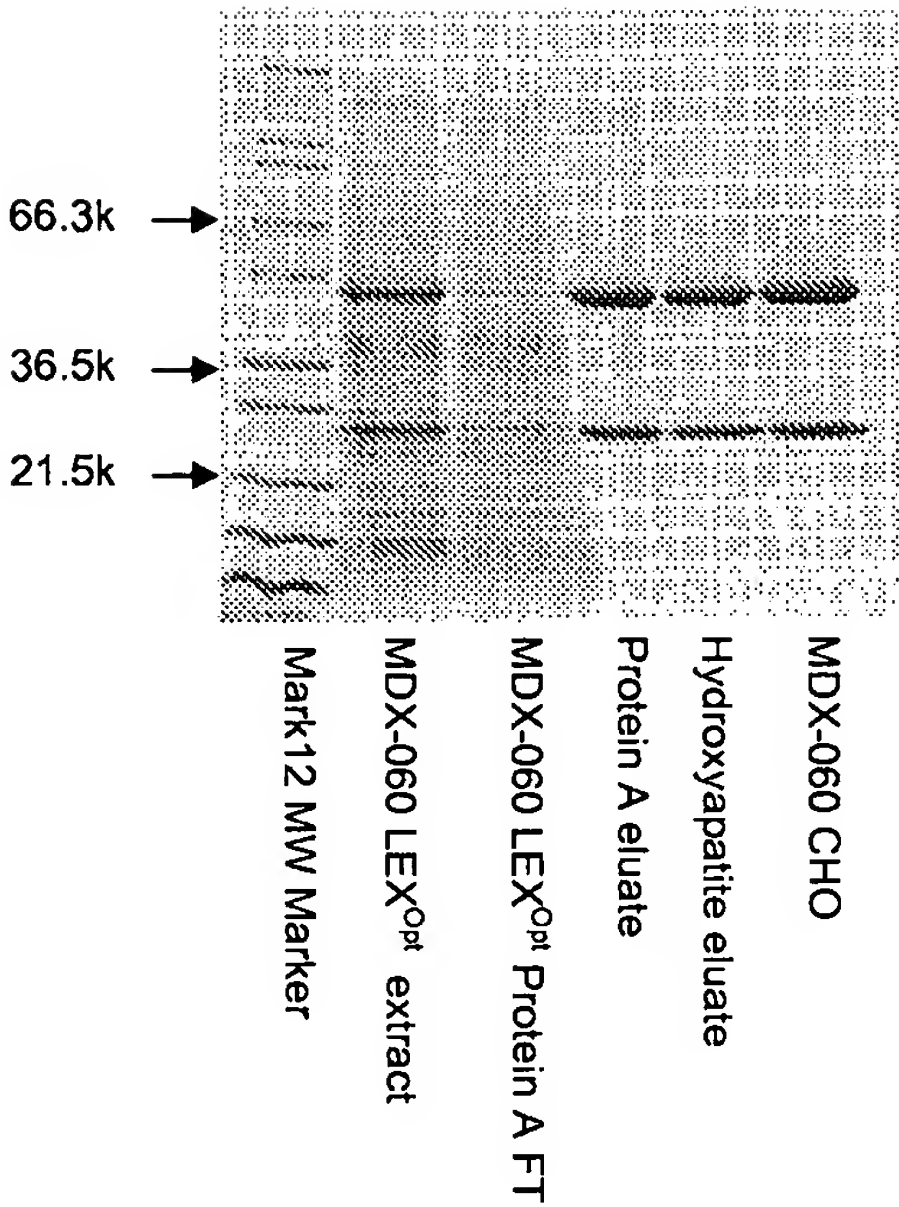


FIG. 41B

47/59

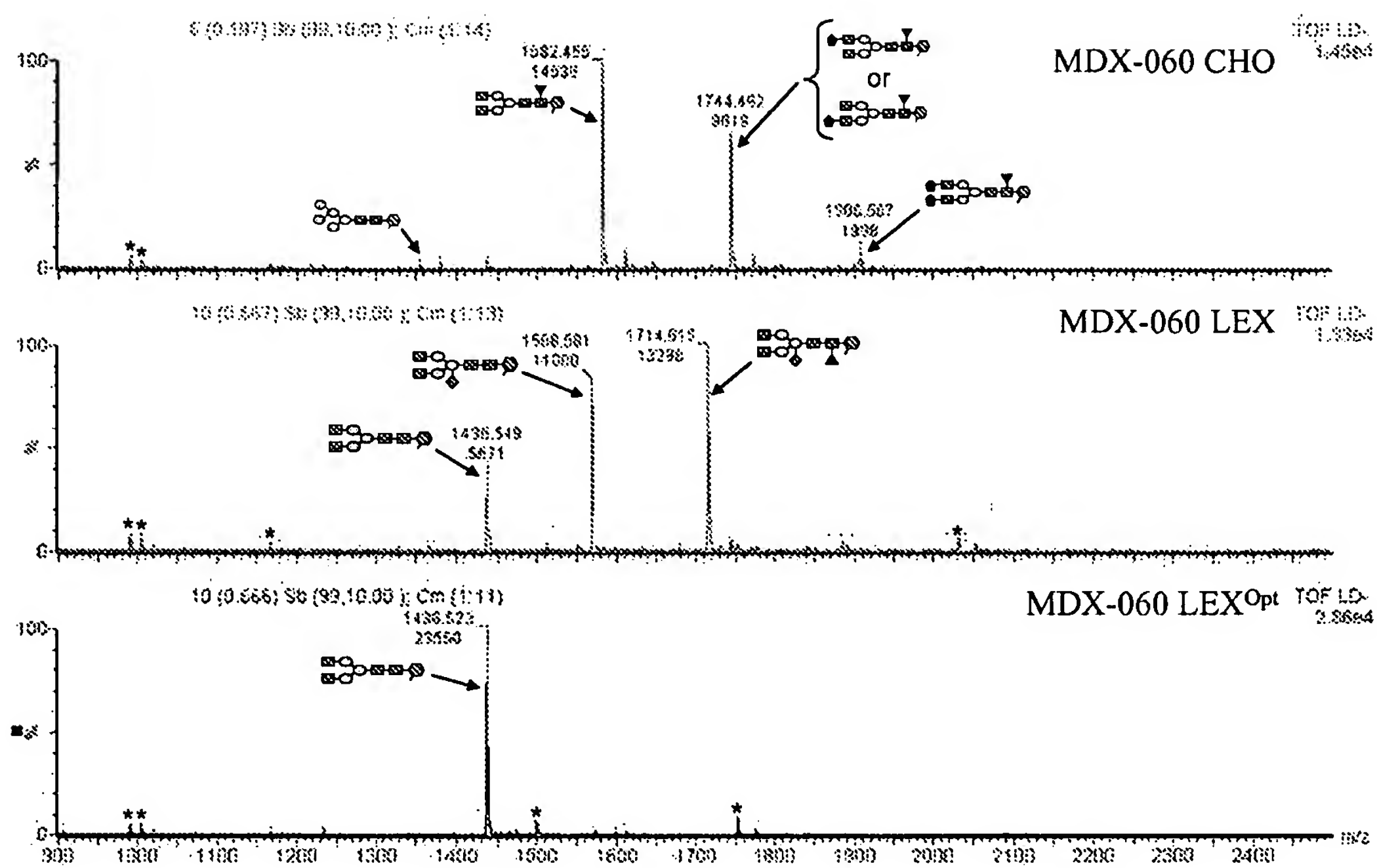


FIG. 42

48/59

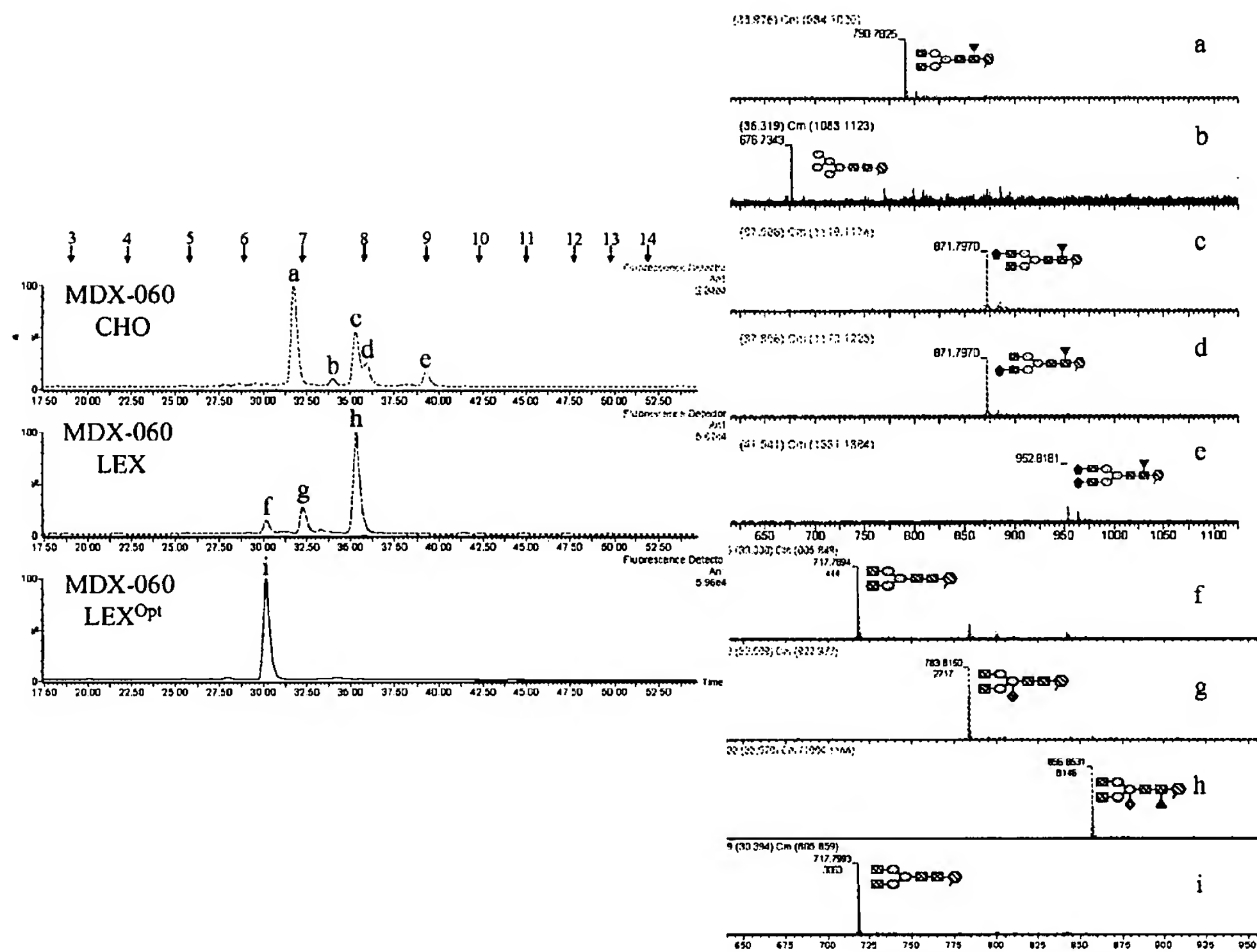


FIG. 43



49/59

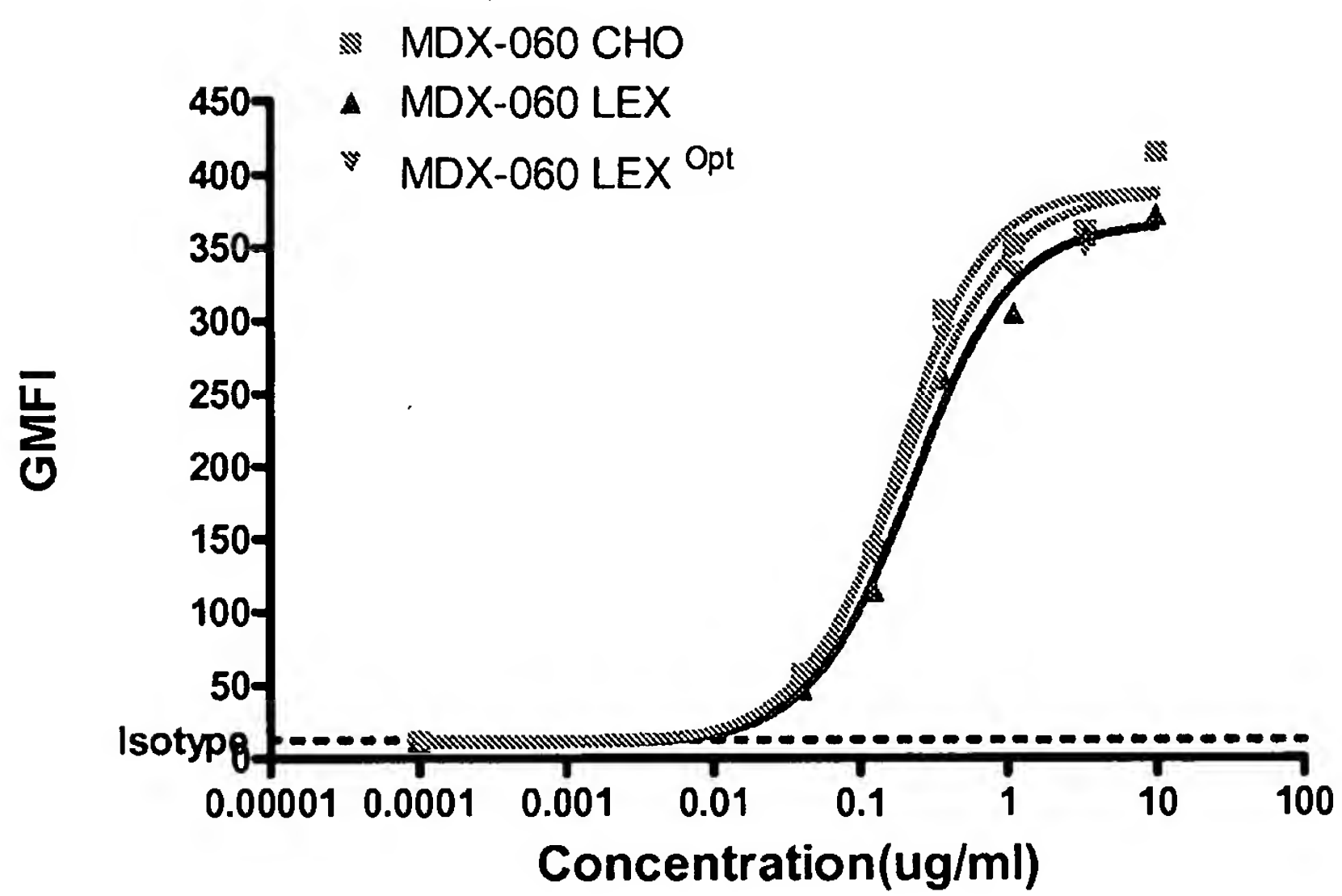
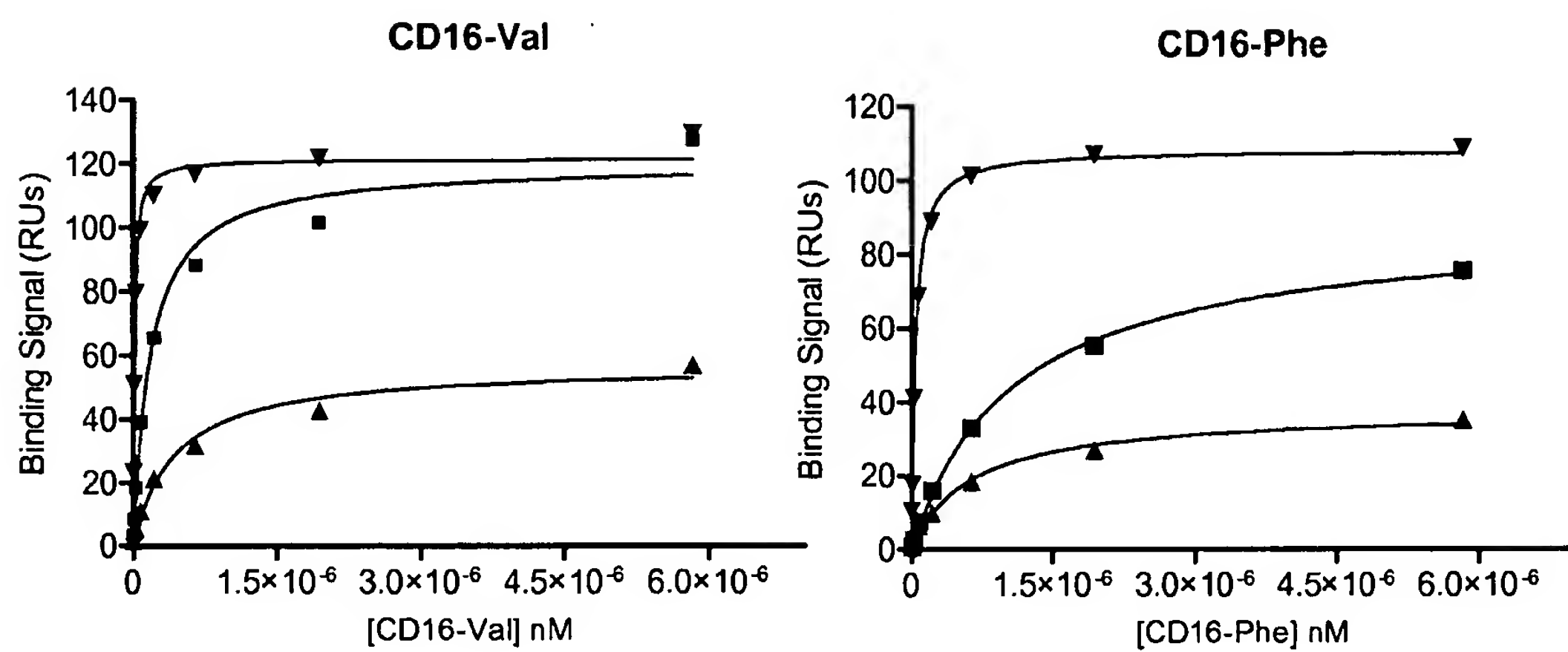


FIG. 44

50/59



	Binding Constants (nM)		
Human FcR	MDX-060 CHO	MDX-060 LEX	MDX-060 LEX <sup>Opt</sup>
CD16-Val	177	427	12
CD16-Phe	1091	641	41

FIG. 45

51/59

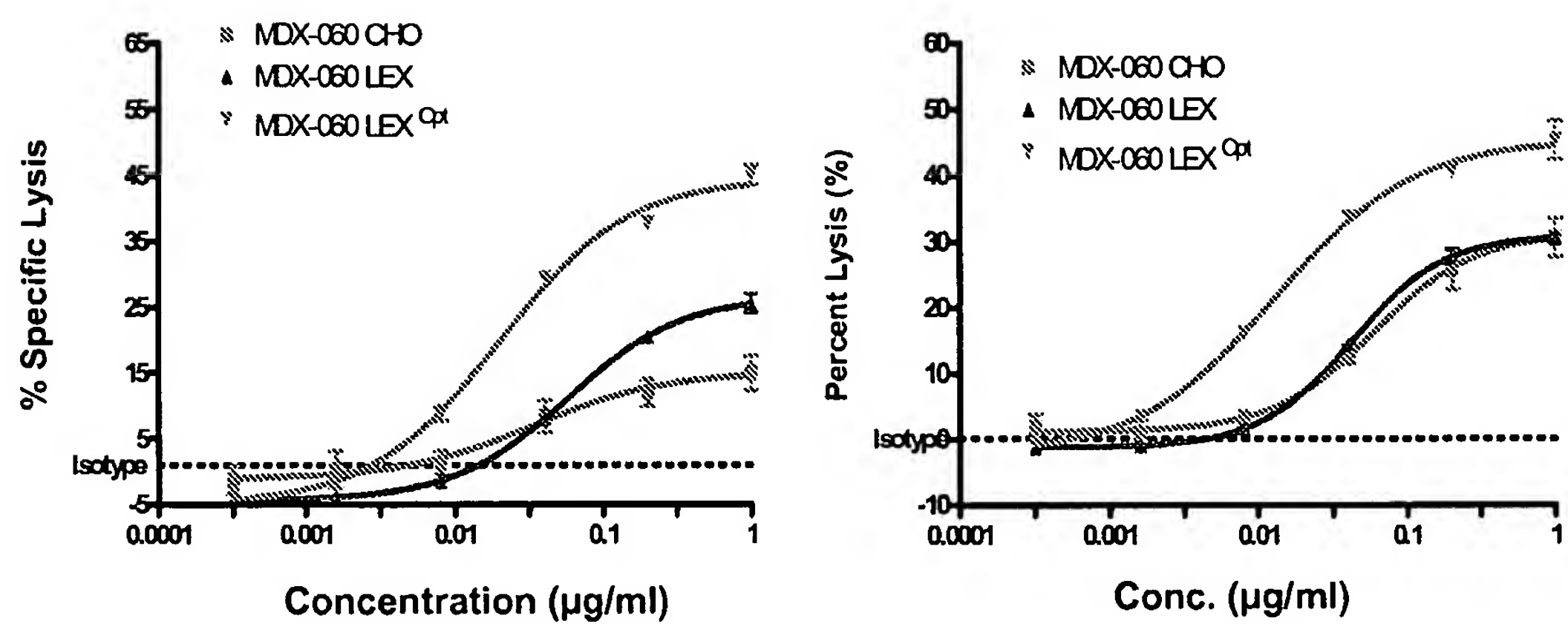


FIG. 46

# MDXA01 Intact Mass Analysis

1: TOF MS ES+  
320

080269\_MDXA01: 280 (10.263) M1 [Ev-113394, 1140] (Gs: 0.750, 2842.3934, 0.10, L33, R33); Sm (SG, 2x4.00); Cm (259.427)

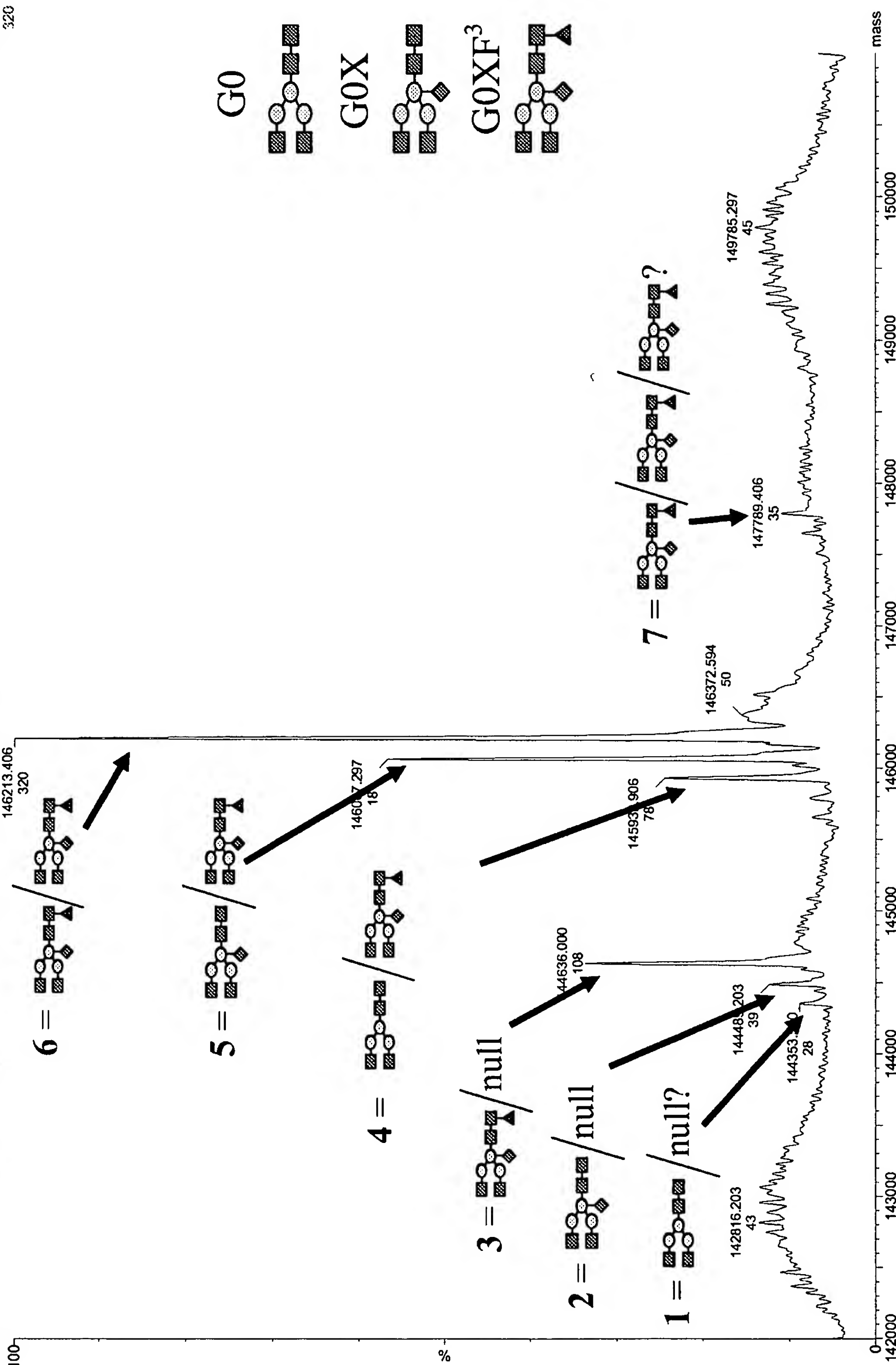


FIG. 47

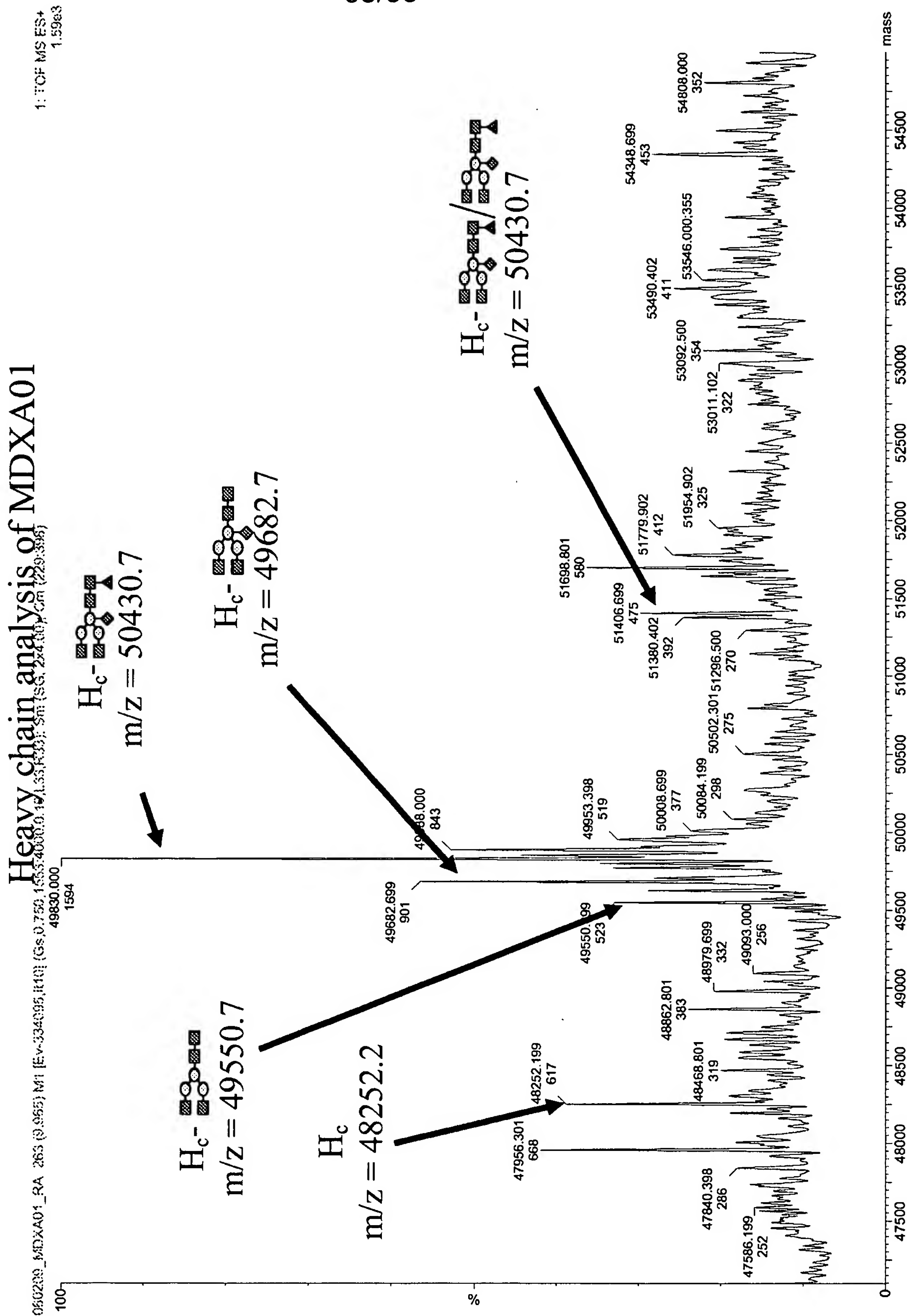


FIG. 48

# MDXA04 Intact Mass Analysis

WO 2007/084926

PCT/US2007/060646

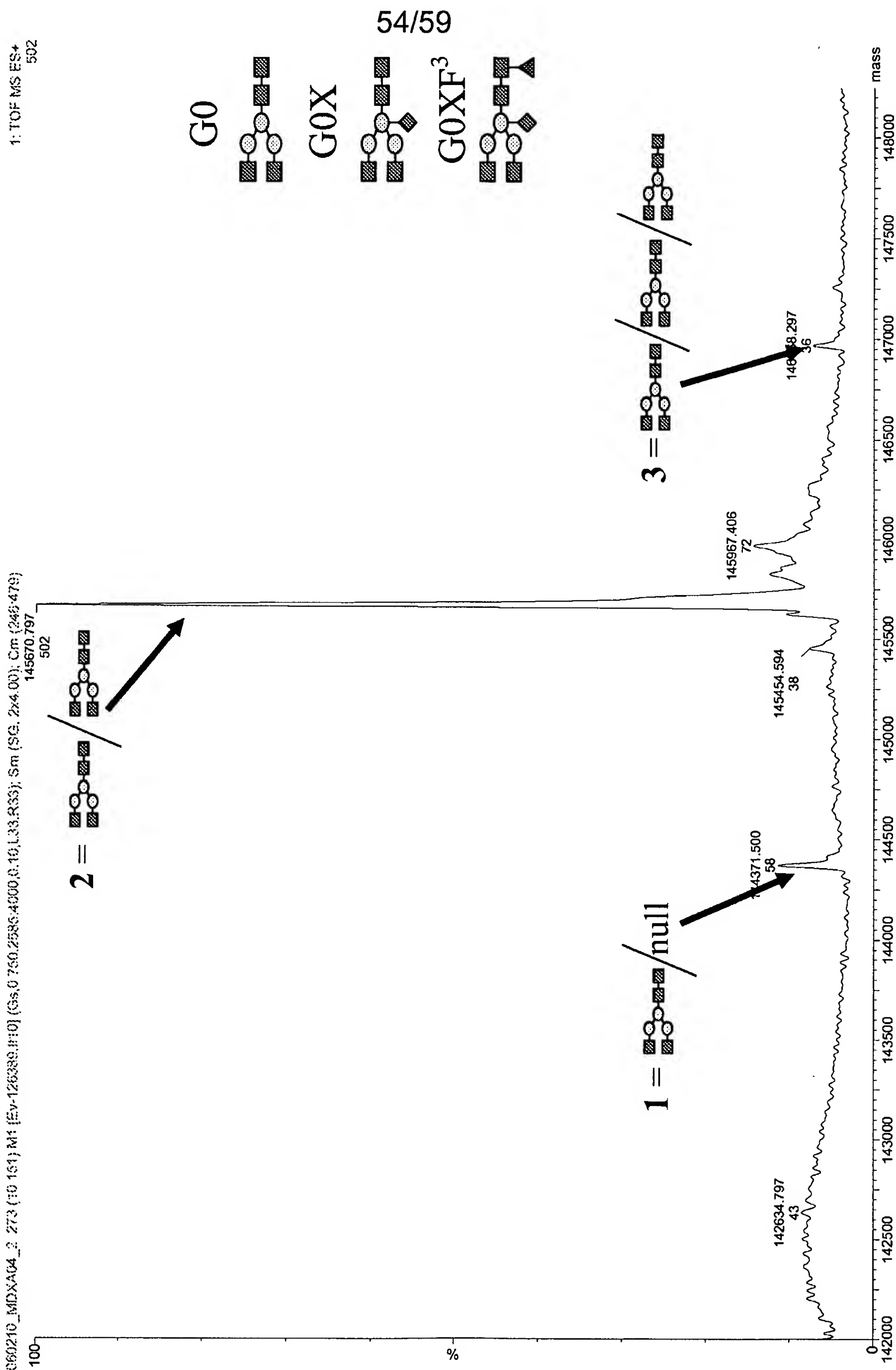


FIG. 49



# Heavy chain analysis of MDXA04

1: TOF MS ES+  
3.13e3

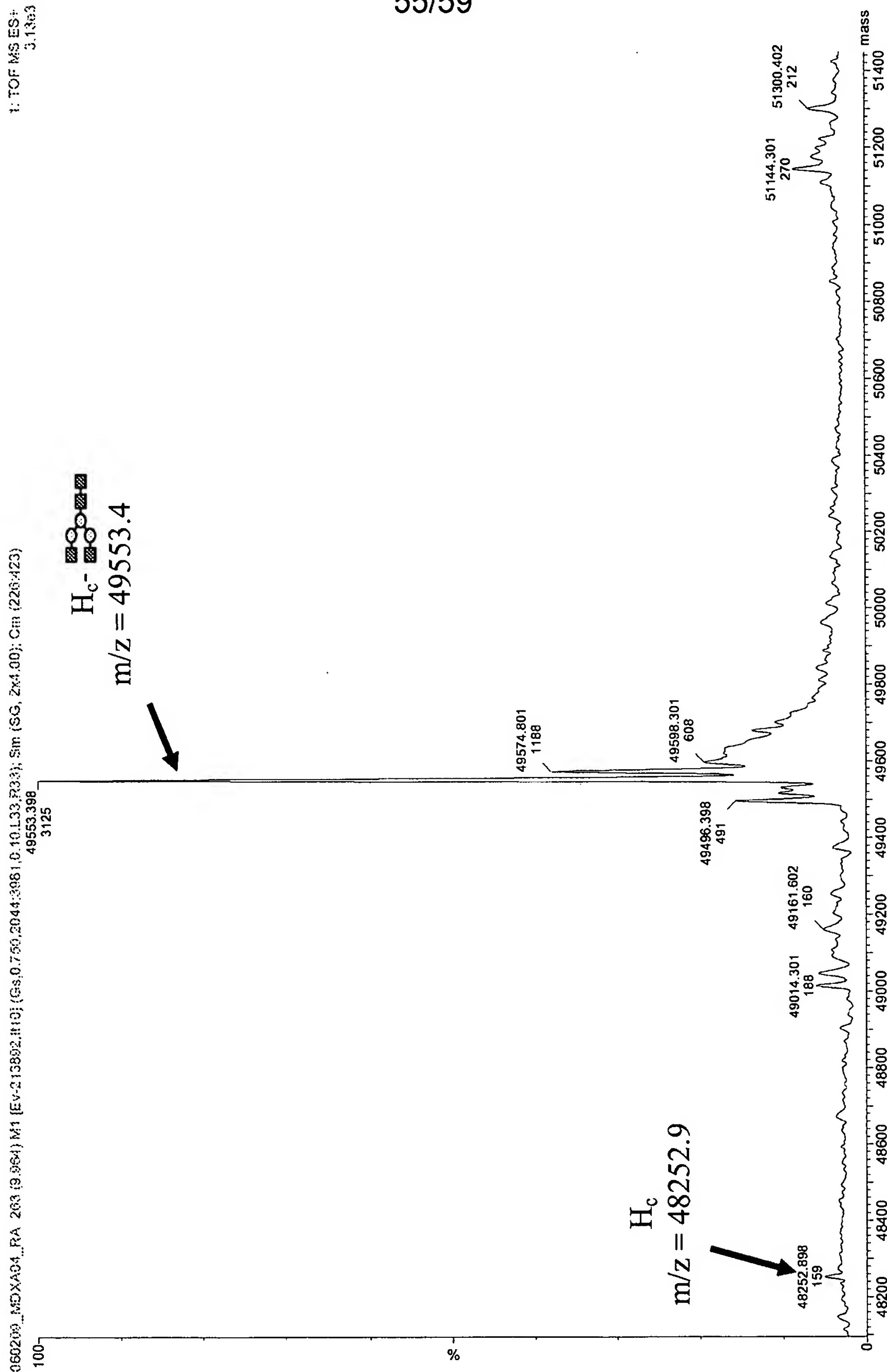


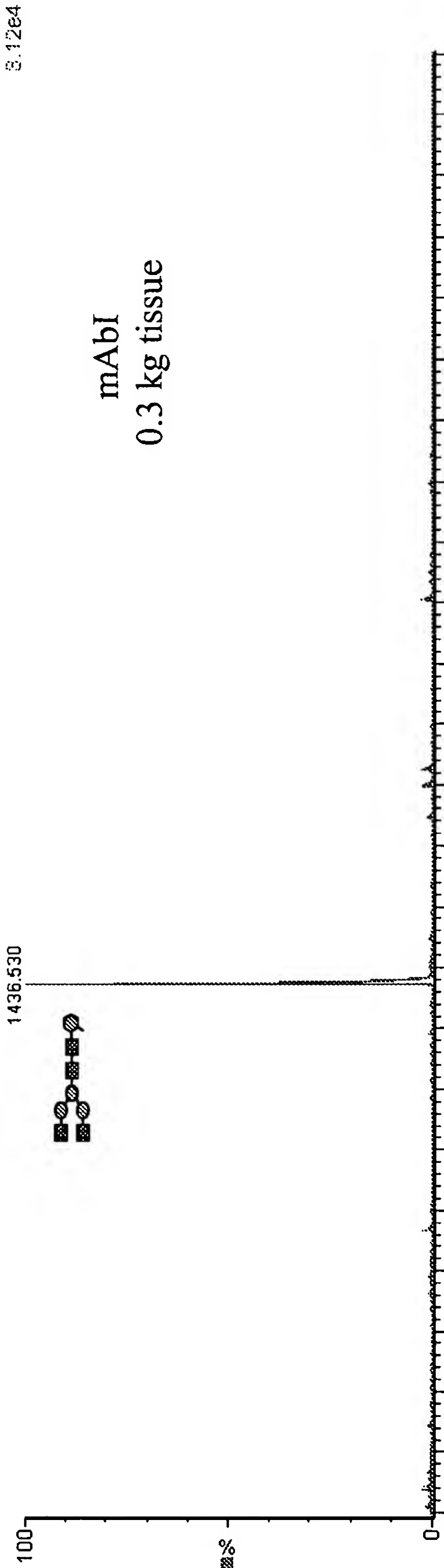
FIG. 50

56/59

# MALDI analysis of N-glycans from scaled-up mAbI04 line

TOF LD-  
3.12e4

mAbI  
0.3 kg tissue



TOF LD-  
2.22e4

mAbI  
6.5 kg tissue

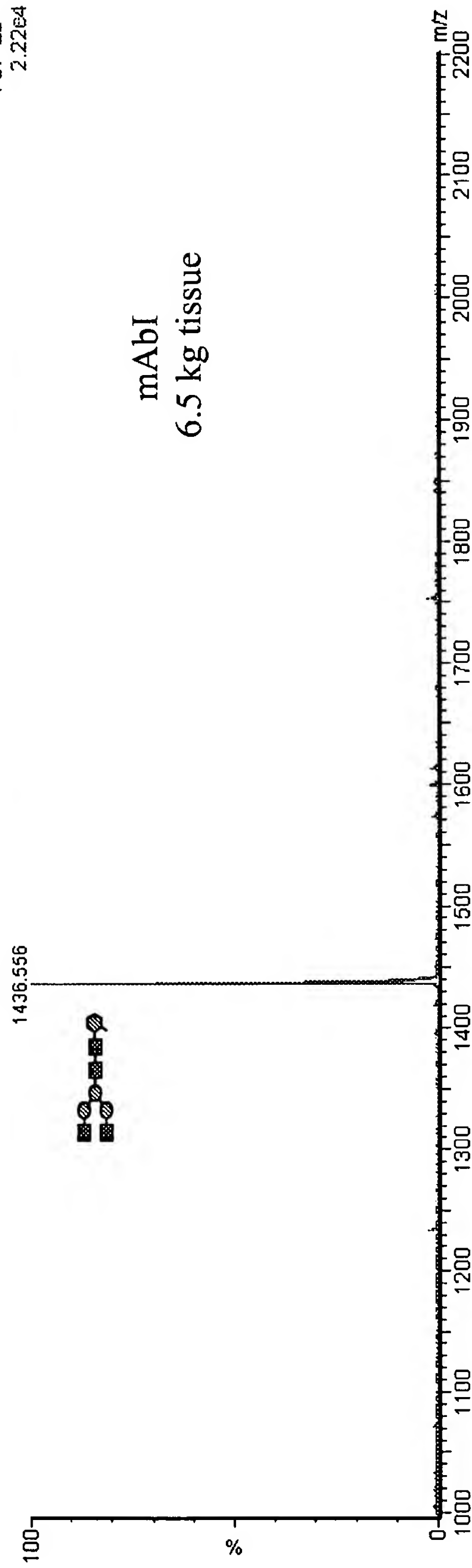


FIG. 51A

HPLC fluorescence analysis of N-glycans from scaled-up mAbI04 line

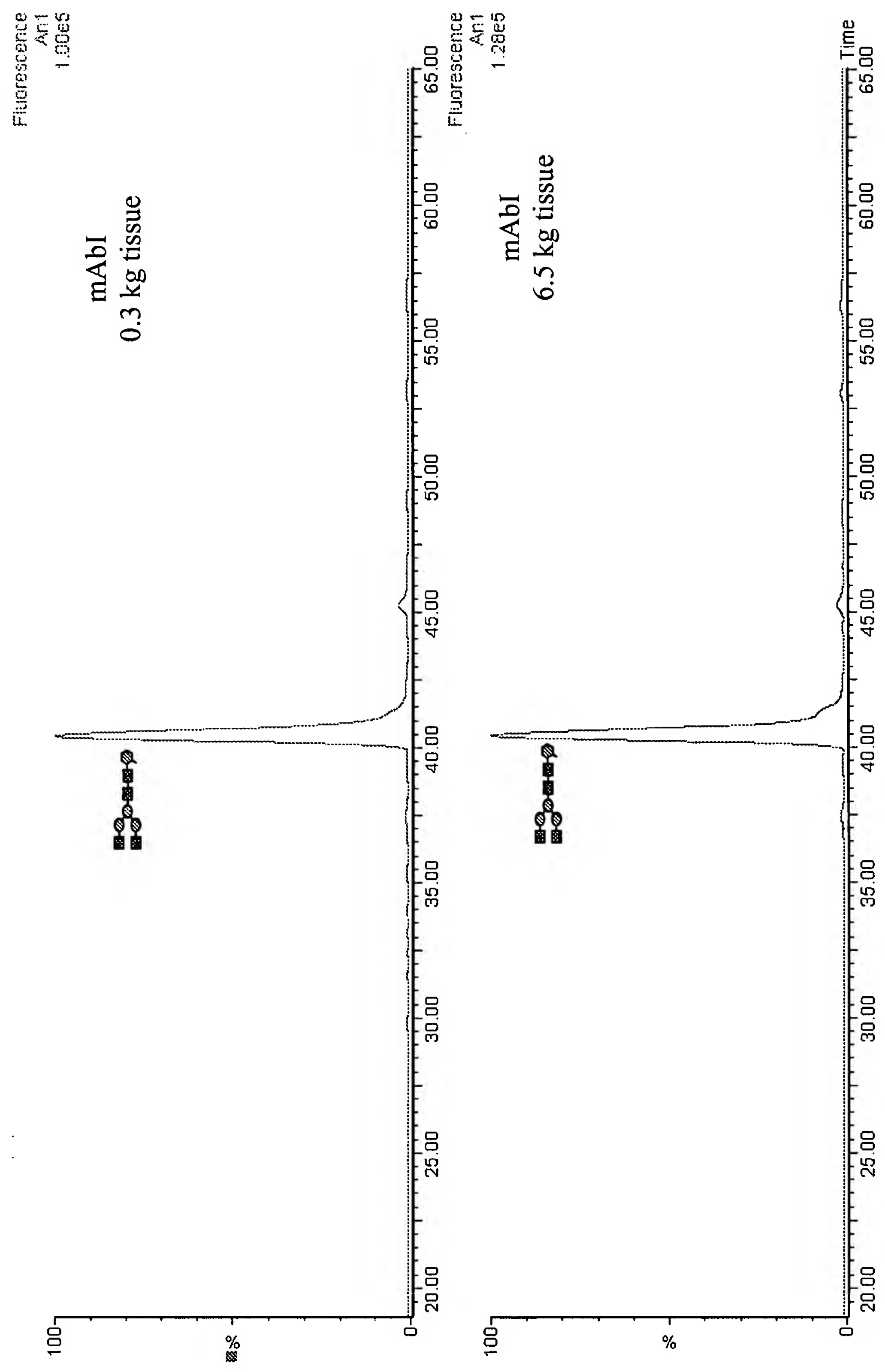


FIG. 51B

## 58/59

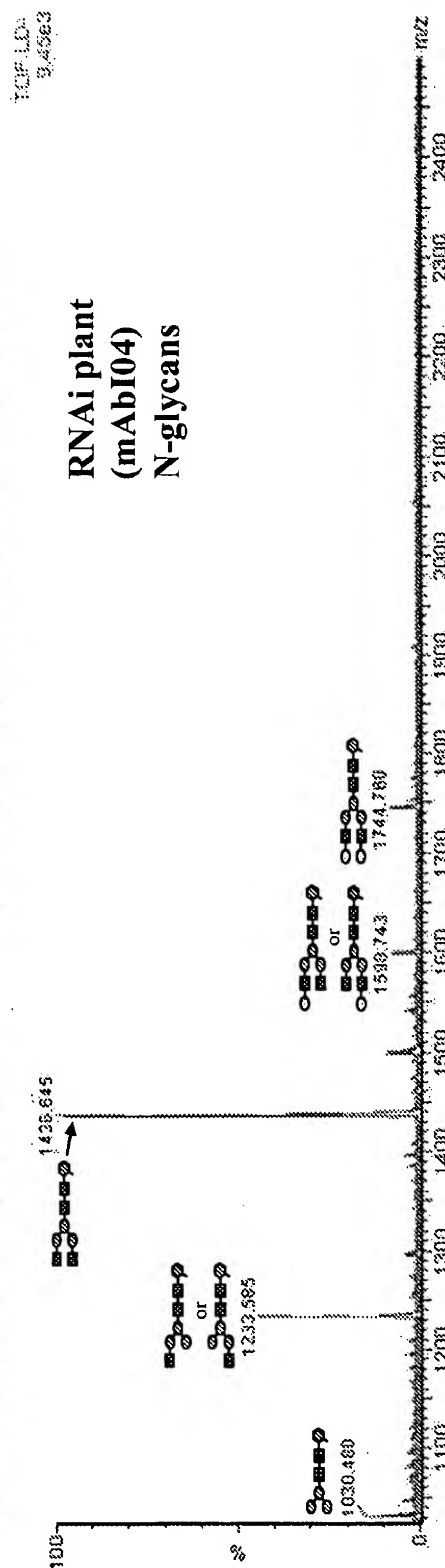
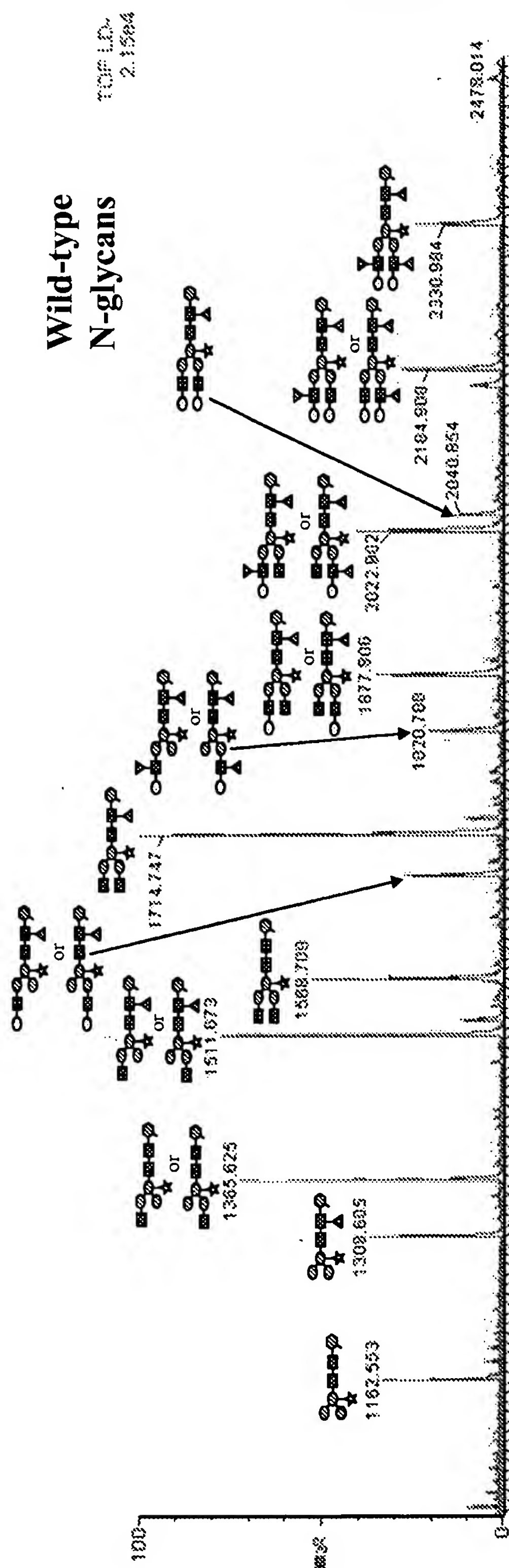


FIG. 52

59/59

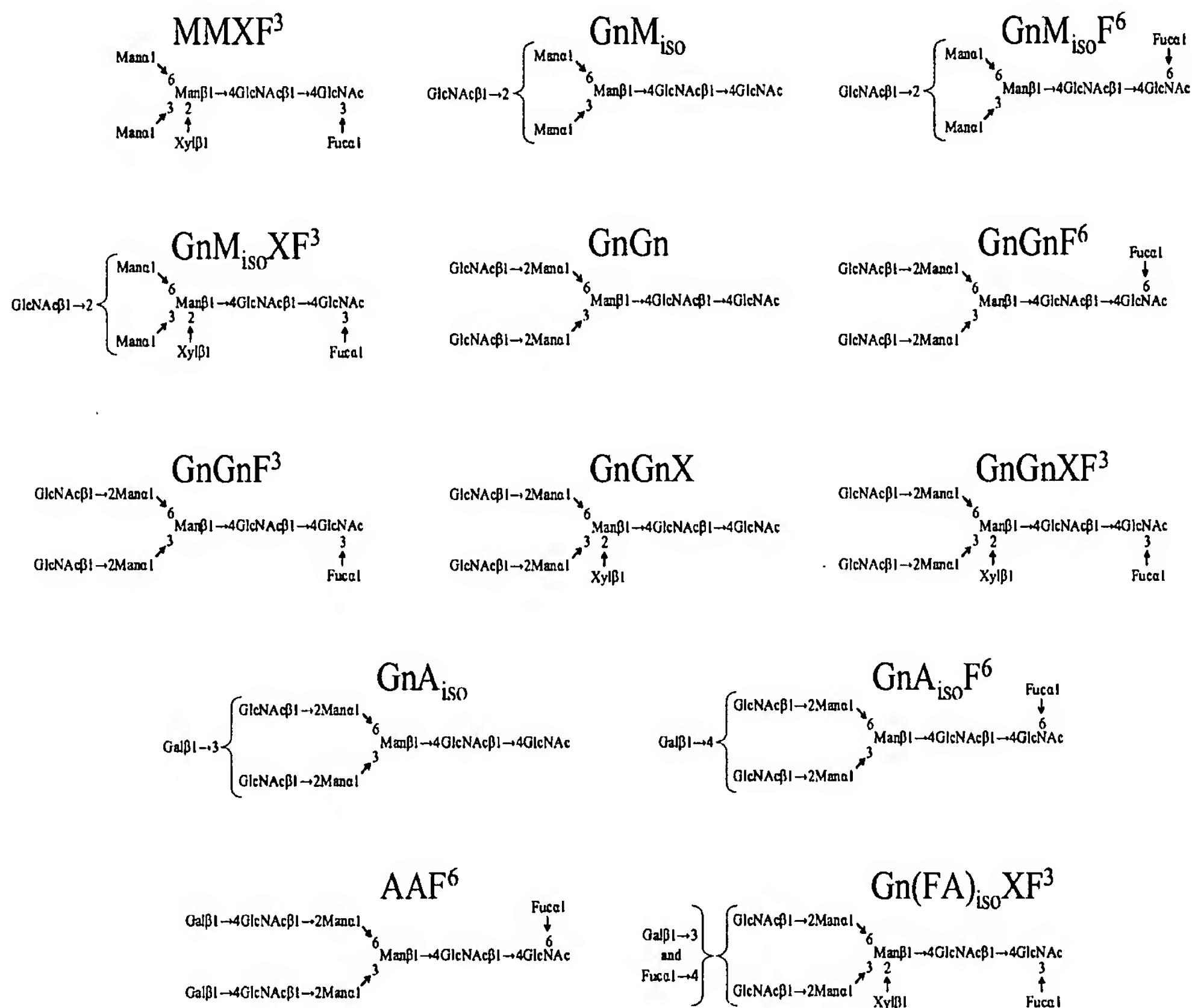


FIG. 53